

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2003, 07:59:32 ; Search time 2512.03 Seconds
(without alignments)
8851.954 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373
Sequence: 1 atgcattgtttagacttga.....gacaaatgttataatg 1373

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 809774376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

EST :
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlun:*
5: em_estlov:*
6: em_estlpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_estl2:*
11: gb_hic:*
12: gb_estl3:*
13: gb_estl4:*
14: gb_estl5:*
15: em_estfun:*
16: em_estcom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_huv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	765.8	55.8	834	13	BI766766 603056866
2	753.2	54.9	948	14	BO707185 AGENCOURT
3	747.2	54.4	963	14	BO671259 AGENCOURT
4	743.4	54.1	777	13	BI819200 603034614
5	725.4	52.8	1071	14	BM921213 AGENCOURT
6	701.4	51.1	731	13	BI871711 603395825

7	688	50.1	828	13	BI596681	603243254
8	677.4	49.3	728 <td>13</td> <td>BI870393</td> <td>603395641</td>	13	BI870393	603395641
9	584.4	42.6	609	13	BI966060	1e72904.x
10	581.4	42.3	1027	17	AF163779	AF163779
11	562	40.9	940	14	BO884231	AGENCOURT
12	534.6	38.9	568	14	BM971606	UI-CF-EC1
13	531.6	38.7	910	12	BG110063	602279667
14	516.8	37.6	531	13	BI824443	603038693
15	493.6	36.0	918	12	BE577781	602092080
16	458.2	33.4	1033	11	AK020909	Mus muscu
17	454.4	33.1	561	10	AW673237	ur70d09.y
18	444	32.3	951	14	BO674188	AGENCOURT
19	442.8	32.3	587	12	BG686319	602638232
20	440.2	32.1	456	13	BI966255	1e72904.y
21	436.8	31.8	440	13	BM128059	1e79b06.y
22	431.6	31.4	538	12	BF821434	MR1-RT003
23	404	29.4	436	17	AO890280	HS_318-B
24	399.4	29.1	785	13	BI762908	603047966
25	394	28.7	413	9	AI422796	tf22e06.x
26	389.2	28.3	698	13	BI906850	603064633
27	374.2	27.3	416	9	AI231866	qmh86c02.x
28	372.6	27.1	584	10	AW917574	EST348878
29	365.4	26.6	422	13	BI677255	1d87a02.x
30	365	26.6	399	9	AI913541	wa12c01.x
31	363.4	26.5	367	17	AO100365	HS_3054.A
32	359	26.1	474	13	BI965174	1d34e07.y
33	358.4	26.1	407	9	AI221985	9q95a10.x
34	357.6	26.0	894	13	BI908274	603068526
35	352.8	25.7	409	10	AW131279	xfl6a04.x
36	347.2	25.3	397	9	AI669243	wc13q01.x
37	342.6	25.0	710	12	BE858778	7f95b06.x
38	341	24.8	377	13	BM509016	1h15b05.y
39	339.8	24.7	372	13	BI677256	1d87a02.y
40	330	24.0	374	13	BM505649	1h15b03.x
41	328	23.9	367	13	BM662622	UI-E-CQ1-
42	327.8	23.8	650	12	BG404836	602420160
43	327	23.8	346	12	BE858822	7g38e04.x
44	327	23.8	471	9	AA221610	my18d09.f
45	326	23.7	337	10	AW195034	xn45g12.x

ALIGNMENTS

RESULT 1
BI766766 834 bp mRNA linear EST 25-SEP-2001
LOCUS 603056866F1 NIH_MGC_122 Homo sapiens CDNA IMAGE:5206217 5',
DEFINITION mRNA sequence.
ACCESSION BI766766
VERSION BI766766.1 GI:15758344
KEYWORDS
SOURCE
ORGANISM human.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LNA11517 row: C column: 18
High quality sequence stop: 772.
Location/Qualifiers

FEATURES
source

1..834

Query Match	Best Local Similarity	55.8%;	Score 765.8;	DB 13;	Length 834;
Matches 818;	Conservative 0;	Mismatches 12;	Indels 5;	Gaps 4;	
272	CCCAGAGAGAGCTGTGTGCACAGAGAGACACAGACCCTGTGGAGACTGATATCCCGACAGACG	331			
1	CCCAGAGAGAGCTGTGTGCAGAGAGAGACAGACCCGTCGGAACTGATATCCCGACAGACG	60			
332	AAGAAAGCCAGAGATCTCTGCGCTTTCCTGTAACCCAGCTAGTTGGCTCGCAGAGATGCAC	391			
61	AAGAAAGCCAGAGATCTCTGCGCTTTCCTGTAACCCAGCTAGTTGGCTCGCAGAGATGCAC	120			
392	CTAAGAGCGCGGAAACAGGGGCTGTGAAGAGGATGCGACGCCATATGAAATTCATCCAC	451			
121	CTAAGAGCGCGGAAACAGGGGCTGTGAAGAGGATGCGACGCCATATGAAATTCATCCAC	180			
452	TGGACACAGAGAGAGACGACGAGCAGAGTGTGAGACGGGACAGTAAAGTGTGCGGAGAGAA	510			
181	GACGCTGACAGAGAGAGAGAGCGAGGACAGTGTGAGACGGGACAGTAAAGTGTGCGGAGAGAA	240			
511	GCCAGAAATCAACAGCTCCAGCCCTCTGCGCTTCAACACCGCCAGATCGGGAGATTATAGTC	570			
241	GCCAGAAATCAACAGCTCCAGCCCTCTGCGCTTCAACACCGCCAGATCGGGAGATTATAGTC	300			
571	ACCGGGCTGGGCTCTACTACTCTGACTGACGTGACACTTTGATGAGGGAGAGCTGTC	630			
301	ACCGGGCTGGGCTCTACTACTCTGACTGACGTGACACTTTGATGAGGGAGAGCTGTC	360			
631	TACCTGAAGCTTGACTTCTGTGTGATGATGTGCTGAGCCCTGCGCTGAGAGAAATTC	690			
361	TACCTGAAGCTTGACTTCTGTGTGATGATGTGCTGAGCCCTGCGCTGAGAGAAATTC	420			
691	TCAGCCACTGGGGCAGTTCCCTCGGGGCCACAGCTCCGCTCTGCCAGATGTCGTGGGCTG	750			
421	TCAGCCACTGGGGCAGTTCCCTCGGGGCCACAGCTCCGCTCTGCCAGATGTCGTGGGCTG	480			
751	TTGGGCTTGGGGCCAGGGTCTCTCCTGTGGGATTCGCGACCCCTCCTGGGCCATCTCAAG	810			
481	TTGGGCTTGGGGCCAGGGTCTCTCCTGTGGGATTCGCGACCCCTCCTGGGCCATCTCAAG	539			
811	GCTGGCCCTTCTCAAGCTACTTGGGACTTTCAGAGTTCAATGAGGGGCCCTGTGCTGCC	870			
540	GCTGGCCCTTCTCAAGCTACTTGGGACTTTCAGAGTTCAATGAGGGGCCCTGTGCTGCC	599			
871	CCACAGTGTCCAGAGCTGCGGGCTCCCTCGACAGCTCTCTGGGACACCGGGTCCCGC-TC	929			
600	CCACAGTGTGTCCAGAGCTGCGGGCTCCCTCGACAGCTCTCTGGGACACCGGGTCCCGC-TC	659			
930	TGCCCCACCTTACGCGGCTCTTTCCTCAAGACTGCGCCCTCCTCTAAGAGCTGCTGGG	989			
660	TGCCCCACCTTACGCGGCTCTTTCCTCAAGACTGCGCCCTCCTCTAAGAGCTGCTGGG	719			
990	CCTGTTCACAGTGTTCATCCCAATATTAATTCAGATATTCGACACTTATCTTACAACCT-1048				
720	CCTGTTCACAGTGTTCATCCCAATATTAATTCAGATATTCGACACTTATCTTAAAAAAT-779				

OY	1049	-CCCCACCGCCACTCTTCACCACAGCAGGCCAACCTGACCCCTTGAAG	1102
Db	780	CCCCCAAGCCCACTCTCACCTCATTAAGTCCCAATTCCTGCACCTTGSAAG	834
RESULT 2			
LOCUS	B0707185		
DEFINITION	B0707185	948 bp	mRNA
ACCESSION	AGENCOURT_8353983 NIH_MGC_113	Homo sapiens	cdna clone IMAGE:6278608
VERSION	BQ707185		
KEYWORDS	BQ707185.1 GI:21846084		
SOURCE	EST.		
ORGANISM	human.		
REFERENCE	Homo sapiens		
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.		
JOURNAL	1 (bases 1 to 948)		
COMMENT	NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgaabs-remail.nih.gov Tissue Procurement: Dr. Mark Watson cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLCM2466 row: n column: 17 High quality sequence start: 24 High quality sequence stop: 550. Location/Qualifiers 1. 948 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:6278608" /clone_1lb="NIH_MGC_113" /lab_host="DH10B (phage-resistant)" /note="Organ: spleen; Vector: pOT8; Site_1: XhoI; Site_2: EcoRI; cDNA made by Oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and NIH-MGC library." NIH-MGC library."		
FEATURES			
source			
BASE COUNT	173 a	334 c	252 g
ORIGIN		187 t	2 others
Query Match	54.9%	Score 753.2:	DB 14: Length 948;
Best Local Similarity	95.6%:	Pred. No. 4.7e-154;	
Matches 807: Conservative	0;	Mismatches 30;	Indels 7; Gaps 3.
OY	305	ACCCGTGCGAATGATGATCCCAGACAGAAGAAGCAGAGATCCTGCGCTTCTGAACC	364
Db	1	ACCCGTGCGAATGATCCCCAGACAAGAANAAGCAGATCTCTGCGCTTCTGAACC	60
OY	365	GACTAGTTGCGCCTCGCAGAAAGTCGACCTAAAGGCGGAAACACGGGCTCGAAGACGA	424
Db	61	GACTAGTTGCGCCTCGAAGAAGTGCACCTAAAGGCGGAAACACGGGCTCGAAGACGA	120
OY	425	TGCGAGCCCATTTATGAAGTTTCATCCAGACCTCGACAGAGACGACGACGAGAGTG	484
Db	121	TGCGAGCCCATTTATGAAGTTTCATCCAGACCTCGACAGAGACGAGAC---GCAGGTGG	176
OY	485	ACGGAGACAGTAGTGCTGCGAGGAAGCCAGATAAACAGAGCTCCAGCCCTTGCCCTACA	544
Db	177	ACGGAGACAGTAGTGCTGCGAGGAAGCCAGATTAACAGAGCTCCAGCCCTTGCCCTACA	236
OY	545	ACCGCAATATGCGGAGATTATAGTACACCGGCGCTGAGCTCTACTACTGTACTGTACAG	604

[illegible]

Query Match	Best Local Similarity	Score	DB	Length
Matches 868; Conservative	91.9%;	Pred. No. 9.5e-153;		
		Mismatches 63; Indels 13; Gaps 7;		
BASE COUNT	194 a	326 c	260 g	183 t
ORIGIN				
Query Match	54.4%;	Score 747.2;	DB 14;	Length 963;
Best Local Similarity	91.9%;	Pred. No. 9.5e-153;		
Matches 868; Conservative	0;	Mismatches 63;	Indels 13;	Gaps 7;
365	AGTTGGGCTCCACAGATGCACTTAAGGCGGAAACGCGCTCCAGAGCGATGCG	428		
1	AGTTGGGCTCCACAGATGCACTTAAGGCGGAAACGCGCGGAAACGCGCTCCAGAGCGATGCG	59		
429	AGCCCATTTATGAAGTTCAATCCAGACCTGAGAGAGCGAGCGGACGCGGATGTGACGG	488		
60	AGCCCATTTATGAAGTTCAATCCAGACCTGAGAGAGCGAGCGGACGCGGATGTGACGG	119		
489	GACAGTGAAGTGGCTGGGAGGAAAGCAAGATCAACAGCTTCAGCCCTTGGCTACAAACGG	548		
120	GACAGTGAAGTGGCTGGGAGGAAAGCAAGATCAACAGCTTCAGCCCTTGGCTACAAACGG	179		
549	CCAGATCGGGGAGTTTATAGTATACCCGGGCTGGCTCTACTACCTGATGTCAGATGCA	608		
180	CCAGATCGGGGAGTTTATAGTATACCCGGGCTGGCTCTACTACCTGATGTCAGATGCA	239		
609	CTTTATGAGGGGAGGAGCTGTACTTGAAGTGGACTGTGCTGAGTGTGCTGGC	668		
240	CTTTATGAGGGGAGGAGCTGTACTTGAAGTGGACTGTGCTGAGTGTGCTGGC	299		
669	CCTGGCTGCTGAGAGAAATTCTACAGCACTCGGCGCACTTCCCTGGGCGCCAGCTCCG	728		
300	CCTGGCTGCTGAGAGAAATTCTACAGCACTCGGCGCACTTCCCTGGGCGCCAGCTCCG	359		
729	CCTGGCGAGGTTGTCGGGCTGTGGGCCCTGGGGCGAGGGTCTCCCTGGGATCGGAC	788		
360	CCTGGCGAGGTTGTCGGGCTGTGGGCCCTGGGGCGAGGGTCTCCCTGGGATCGGAC	419		
789	CCCTCCCTGAGGCGCATCTCAAGGCTGCGCCCTCTCTCACTCACTTGTGAGCTTTCAGAGT	848		
420	CCCTCCCTGAGGCGCATCTCAAGGCTGCGCCCTCTCTCACTCACTTGTGAGCTTTCAGAGT	479		
849	TCACAGAGGGGCTGTGCTTCCACACAGTGTGTCACAGGCTGCGGCTCCCTTCGACAGCT	908		
480	TCACAGAGGGGCTGTGCTTCCACACAGTGTGTCACAGGCTGCGGCTCCCTTCGACAGCT	539		
909	CTCTGGGCGAGCGGCTGCGGCTGTGGCCACCTCAGCGGCTCTTGGTCCAGAGCTGGCCC	968		
540	CTCTGGGCGAGCGGCTGCGGCTGTGGCCACCTCAGCGGCTCTTGGTCCAGAGCTGGCCC	599		
969	TCCTCTTAGAGCTGCTGGGCTGTTCACGCTGTTTTCATCCATCCCATTAATACAGTATT	1028		
600	TCCTCTTAGAGCTGCTGGGCTGTTCACGCTGTTTTCATCCATCCCATTAATACAGTATT	659		
1029	CCCACTTTATTTTACAACTCCCCC- ACCGCCACTCTTCACCTCACTAGTACCTCCCAATC	1087		
660	CCCACTTTATTTTACAACTCCCCC- ACCGCCACTCTTCACCTCACTAGTACCTCCCAATC	719		
1088	CC- ----TGACCTTTGAGGCGCCCAAGTATCTTCGACTCCCGCTGGCCACAAGCCCGAG	1143		
720	CCCTGACCCCTTTGAGGCGCCCAAGTATCTTCGACTCCCGCTGGCCACAAGCCCGCA	779		

QY	1144	GGCAATT-CTGTTCACCTGACCTCTGTGGGCAAGG--ATGGGCTCAGAAAGCCCACTTCAGG	1201
Db	780	GGCAATTGGGGTTCTACTGAACTCTGTGGGGGAAGGAAGGGGCTCCAAAAGACCCCTTTGGG	839
QY	1202	CACATAAGAGGGGCT--GGACCTGGCGGCGAGAAAGCCAAAGAGACTGGG---CCTAGGCCA	1256
Db	840	ACTTAAGAGGGGCTTGACACTGTGGGGGAAGAAACCCAAAAAATGTGGGGCCCAAGGCCCA	899
QY	1257	GGAGTTTCCCAATGTGAGGGCGAGAAACAAGACAGCTCTCC	1300
Db	900	GAATTTCCCAAGAGGGGCGGGAAGAAAAACAAAAAGCTCC	943
RESULT 4			
LOCUS	Bi1819200	777 bp	mRNA
DEFINITION	603034614F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175698 5',		linear EST 04-OCT-2001
ACCESSION	Bi1819200		
VERSION	Bi1819200.1	GI:15930750	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;		
AUTHORS	Mammalia: Eutheria: Primates: Catarrhini: Homnidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LNL at:		
	http://image.lnl.gov		
	Plate: LHAM11437 row: 1 column: 03		
	High quality sequence stop: 759.		
FEATURES			
source	1..777		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5175698"		
	/clone_id="NIH_MGC_115"		
	/lab_host="DH10B"		
	/note="Organ: pooled brain, lung, testis; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Grueter (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC library."		
BASE COUNT	140 a 268 c 215 g 154 t		
ORIGIN			
	Query Match 54.1%; Score 743.4; DB 13; Length 777;		
	Best Local Similarity 99.1%; Prid. No. 6.2e-152;		
	Matches 758; Conservative 0; Mismatches 6; Indels 1; Gaps 1;		
QY	288	GGCAGAGGAGGACCGCGTGGCACTGTAATCCAGACAGAAAGCCAGATCC	347
Db	13	GGCAGAGGAGGACCGCGTGGCACTGTAATCCAGACAGAAAGCCAGATCC	72
QY	348	TGCGGCTTTCTGAAACGACTAGTTGGGCTTCGCAAGAGTGACCTAAAGCGGAAAAAC	407
Db	73	TGCGGCTTTCTGAAACGACTAGTTGGGCTTCGCAAGAGTGACCTAAAGCGGAAAAAC	132
QY	408	ACGGGCTCGAAAGCGATGCGAGCCCATTTAGAGTTATCCAGACCTGACAGAGCGG	467

Db	133	ACGGGCTCGAAGACGAGATCGCAGCCCATTTATGAAGTTATCTCCAGCACTGGACAGGAGG	192
QY	468	AGGCCAGGACAGGTGTGGACGGGACAGTGAAGTGGCTGGGAGAGAGCCAGATCAACAGCTC	527
Db	193	AGGCCAGGACAGGTGTGGACGGGACAGTGAAGTGGCTGGGAGAGAGCCAGATCAACAGCTC	252
QY	528	CAGCCCTCTGGGCTACACAACCCGCGAGATCGGAGTTTATCTACACCCGGGCTGGGCTCTA	587
Db	253	CAGCCCTCTGGGCTACACAACCCGCGAGATCGGAGTTTATCTACACCCGGGCTGGGCTCTA	312
QY	588	CTACCTGTACGTGATAGGAGCACTTTGATGAGGGAGAGGCTCTTACCTGAAGCTTGACTT	647
Db	313	CTACCTGTACGTGATAGGAGCACTTTGATGAGGGAGAGGCTCTTACCTGAAGCTTGACTT	372
QY	648	GCTGGTGGATGATGTGCTGAGCCCTGCGCTGCGCTGGAGAGAAATTCAGCCACTGCGGCGAG	707
Db	373	GCTGGTGGATGATGTGCTGAGCCCTGCGCTGCGCTGGAGAGAAATTCAGCCACTGCGGCGAG	432
QY	708	TTCCCTCGGGGCCCCACAGCTCGCGCTCTGCGCAGAGTGTCTGGGCTGTTGGGCTCGCGGCAAG	767
Db	433	TTCCCTCGGGGCCCCACAGCTCGCGCTCTGCGCAGAGTGTCTGGGCTGTTGGGCTCGCGGCAAG	492
QY	768	GTCCTCCCTGGGGAATCCGACACCTCCCTCGGAGCCCATCTCAAGGCTGCGCCCTTCTCTAC	827
Db	493	GTCCTCCCTGGGGAATCCGACACCTCCCTCGGAGCCCATCTCAAGGCTGCGCCCTTCTCTAC	552
QY	828	CTACTCGGACTCTTCCAGGTTTCACTGAGGGGACCCTGCTGCCACCAAGTCGATCCAGGC	887
Db	553	CTACTCGGACTCTTCCAGGTTTCACTGAGGGGACCCTGCTGCCACCAAGTCGATCCAGGC	612
QY	888	TGCGGGCTCCCTCGACACGCTCTCTGGGCAACCCGCTCCCTCTCTGCCCCACCCCTCAGCCGC	947
Db	613	TGCGGGCTCCCTCGACACGCTCTCTGGGCAACCCGCTCCCTCTCTGCCCCACCCCTCAGCCGC	672
QY	948	TCTTTGCTCGAGACTGCGCCCTCCCTCTAGAGAGGCTGCGCTGGGCGCTGTCAGCTGTTTTC	1007
Db	673	TCTTTGCTCGAGACTGCGCCCTCCCTCTAGAGAGGCTGCGCTGGGCGCTGTTTTC	732
QY	1008	ATCCACATTAATATACAGTATTCGCCACTCTTATCTTACACTGCC	1052
Db	733	ATCCACATTAATATACAGTATTCGCCACTCTTATCTTACACTGCC	776
RESULT 5			
LOCUS	BM921213	1071 bp	mRNA linear EST 12-MAR-2002
DEFINITION	AGENCOURT 6633046 NIH_MGC_115 Homo sapiens CDNA clone IMAGE:5752561		
ACCESSION	BM921213		
VERSION	BM921213.1	GI:19371592	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	1 (bases 1 to 1071)		
JOURNAL	NIH-MGC http://mgc.ncl.nih.gov/ .		
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC)		
	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: gsa@femail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Agencourt Bioscience Corporation		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAM12786 row: P column: 02		
	High quality sequence stop: 656.		
FEATURES	Location/Qualifiers		
SOURCE	1..1071		

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5752561"
/clone_1lb="NIH_MGC_115"
/lab_host="DH10B"
/Note="Organ: pooled brain, lung, testis; Vector:
pcmv-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

BASE COUNT      226 a      346 c      279 g      220 t
ORIGIN

Query Match      52.8%; Score 725.4; DB 14; Length 1071;
Best Local Similarity 93.4%; Pred. No. 5.6e-148;
Matches 836; Conservative 0; Mismatches 46; Indels 13; Gaps 7;

OY 412 GCTGGAAGAGGATCGACAGCCATTATGATCATCCAGACCTGACAGAGGAGCG 471
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 12 GCTGGAAGAGGATCGACAGCCATTATGATCATCCAGACCTGACAGAGGAGCG 71

OY 472 CAGGAGGTGGAGCGGAGACAGTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGG 531
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 72 CAGGAGGTGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 131

OY 532 CCTGTGGGCTACAAACCGGACAGATCGGGAGATTATAGTACACCGGGCTGTACTAC 591
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 132 CCTGTGGGCTACAAACCGGACAGATCGGGAGATTATAGTACACCGGGCTGTACTAC 191

OY 592 CTGTACTGTCAAGGTGCACTTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 651
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 192 CTGTACTGTCAAGGTGCACTTTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGG 251

OY 652 GTGATGATGTGTGGGCGGCTGGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 711
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 252 GTGATGATGTGTGGGCGGCTGGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 311

OY 712 CTGCGGGCCCGACAGTCCGCTGTGCGAGGTGCTGGGCTGTGGGCTGTGGGCTGT 771
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 312 CTGCGGGCCCGACAGTCCGCTGTGCGAGGTGCTGGGCTGTGGGCTGTGGGCTGT 371

OY 772 TCCCTGGGAGATCGGACCTCCCTCGGCGCATCTCAAGGCTGCCCTCTCTCTAC 831
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 372 TCCCTGGGAGATCGGACCTCCCTCGGCGCATCTCAAGGCTGCCCTCTCTCTAC 431

OY 832 TTGCGAGCTTTCAGAGTCACTGAGAGGCGCTGGTCTCCCAAGTCTCCAGGCTGCC 891
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 432 TTGCGAGCTTTCAGAGTCACTGAGAGGCGCTGGTCTCCCAAGTCTCCAGGCTGCC 491

OY 892 GGGTCCCTCGACAGCTCTGTGGGACCGGCTCCCTGTGCGCCACCTCAGCGCTCTT 951
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 492 GGGTCCCTCGACAGCTCTGTGGGACCGGCTCCCTGTGCGCCACCTCAGCGCTCTT 551

OY 952 TGTCTCAGACCTGCGCTCTCTTCTAGAGGCTGCTGGGCTGTTCACGTCTTCATCC 1011
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 552 TGTCTCAGACCTGCGCTCTCTTCTAGAGGCTGCTGGGCTGTTCACGTCTTCATCC 611

OY 1012 CACGTAATATAGTAATTCCTCCACTCTTATCTTACAACCTCCCGACGCTCTCCACT 1071
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 612 CACGTAATATAGTAATTCCTCCACTCTTATCTTACAACCTCCCGACGCTCTCCACT 671

OY 1072 CACTAGCTCCCAATCCCTGTGAGGCGCCCACTGATCTGCACT -CCGCCCTGGC 1130
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 672 CACTAGCTCCCAATCCCTGTGAGGCGCCCACTGATCTGCACTCCGCCCTGGC 731

OY 1131 CACAGAGCCCG -AGGCAATGTG -TTCAGTGTACTCTGTGGGCAAGA -TGGGTCCAGA 1186
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 732 CACAGAGCCCGCAAGGCAATGTGTGTTCAGTGTGCTGGGCAAGAATGGGTCCAGA 791

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5405478"
/clone_1lb="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/Note="Organ: liver; Vector: pcmv-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

BASE COUNT      125 a      242 c      228 g      136 t
ORIGIN

Query Match      51.1%; Score 701.4; DB 13; Length 731;
Best Local Similarity 99.3%; Pred. No. 9e-143;
Matches 726; Conservative 0; Mismatches 1; Indels 4; Gaps 2;

OY 132 CCTGGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 251
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 1 CTTGGCGCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60

OY 252 GCTGTCCGCGC--CAGGAGCGGCGCCAGAGAGCTGTGGGAGAGAGGAGGAGGAGCC 308
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 61 GCTGTCCGCGCAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGCC 120

OY 309 GTGGAATGTAATCCCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 368
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 121 GTGGAATGTAATCCCGACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180

OY 369 AGTTGGGCTGTGAGAGTGCATTAAGGCGGAGAAACAGGGGCTGCAAGAGCATGCC 428
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 181 AGTTGGGCTGTGAGAGTGCATTAAGGCGGAGAAACAGGGGCTGCAAGAGCATGCC 240

```

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
Oy 429	AGCCATTATGAGTTCATCCACGACCTGGACAGGACGGAGCGACGGAGCTGTGACGG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Db 241	AGCCATTATGAGTTCATCCACGACCTGGACAGGACGGAGCGACGGAGCTGTGACGG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Oy 489	GACACTGATGAGCTGTGGAGGAGCAGATCAACAGCTCCAGCCCTTGCCTACACCG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Db 301	GACACTGATGAGCTGTGGAGGAGCAGATCAACAGCTCCAGCCCTTGCCTACACCG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Oy 549	CCAGATCGGGAGGATTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGTACAGTGA	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Db 361	CCAGATCGGGAGGATTATAGTCACCCGGGCTGGGCTCTACTACCTGTACTGTACAGTGA	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Oy 609	CTTTGATGAGGGAAGGCTGTCTTACCTCAAGCTGACCTGTGCTGTGATGTGTCTGTGC	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Db 421	CTTTGATGAGGGAAGGCTGTCTTACCTCAAGCTGACCTGTGCTGTGATGTGTCTGTGC	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Oy 669	CCTGGGCTGCTCGAGAGGAATTTCTACGCCACTGTGGGCTCCCTCGGGGCCCACTCGG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Db 481	CCTGGGCTGCTCGAGAGGAATTTCTACGCCACTGTGGGCTCCCTCGGGGCCCACTCGG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Oy 729	CCTTCGCCAGGATGATGTGGAGCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Db 541	CCTTCGCCAGGATGATGTGGAGCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Oy 789	CCTTCGCCAGGATGATGTGGAGCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Db 600	CCTTCGCCAGGATGATGTGGAGCTGTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Oy 849	TCACAGAGGGGGCCGCTGCTCCGACAGCTGCTCCGACAGGCTGCTCCGACAGGCTGCTCCG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Db 660	TCACAGAGGGGGCCGCTGCTCCGACAGCTGCTCCGACAGGCTGCTCCGACAGGCTGCTCCG	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Oy 909	CTCTGGGACAC 919	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			
Db 720	CTCTGGGACAC 730	B1596681	603243254F1	NIH_MGC_96	Homo sapiens	CDNA clone	IMAGE:5285892	5'			

Query Match	Best Local Similarity	Score	DB	Length
Matches 732; Conservative	95.3%;	688;	13;	828;
	0;	7.8e-140;		
	Mismatches 30;	Indels 6;	Gaps 2;	
<p>BASE COUNT 155 a 278 c 223 g 172 t</p> <p>ORIGIN</p>				
OY	312	GGAAGTGAATCCCGCAGACAGAGAAAAGCCAGATCTCTGGCCTTTCTGTAACCGACTAGT	371	
DB	43	GGAACTTGAAATCCCGCAGACAGAGAAAAGCCAGATCTCTGGCCTTTCTGTAACCGACTAGT	102	
OY	372	TGCGGCTCGCAGAACTGCGACCTTAAAGGCGGAAACACGCGGCTCGAAGAGCATGCGAGC	431	
DB	103	TGCGGCTCGCAGAACTGCGACCTTAAAGGCGGAAACACGCGGCTCGAAGAGCATGCGAGC	162	
OY	432	CCATTATGAAGTTCAATCCACGACTGCGACAGAGGAGCGCAGGAGAGTGTGGACGGGAC	491	
DB	163	CCATTATGAAGTTCAATCCACGACTGCGACAGAGGAGCGCAGGAGAGTGTGGACGGGAC	222	
OY	492	AGTGAATGAGCTGGGAGGAGAAAGCAGAAATCAACAGCTCCAGCCCTTGCGGTATCAACCGCA	551	
DB	223	ATTGATGAGCTGGGAGGAGAAAGCAGAAATCAACAGCTCCAGCCCTTGCGGTATCAACCGCA	282	
OY	552	GATCGGGAGATTATAGTACACCGGGGCTGGGCTTACTACTCTACTGTCAAGTGCACATT	611	
DB	283	GATCGGGAGATTATAGTACACCGGGGCTGGGCTTACTACTCTACTGTCAAGTGCACATT	342	
OY	612	TGATGAGGGGAGAGCGTGTCTACTGAAGTGGAGCTTGGGAGGAGTGGTGTCTGGCCCT	671	
DB	343	TGATGAGGGGAGAGCGTGTCTACTGAAGTGGAGCTTGGGAGGAGTGGTGTCTGGCCCT	402	
OY	672	GCGCTGCCTGGAGGAATTTCTCAGCCACTGCGGCCAGGTTCCCTCGGGGCCCCAGCTCGGCT	731	
DB	403	GCGCTGCCTGGAGGAATTTCTCAGCCACTGCGGCCAGGTTCCCTCGGGGCCCCAGCTCGGCT	462	
OY	732	CTGCCAGAGTGTGTGGGCTGTGTGGCCCTGCGGCCAGAGGTCTCTCCCTGGGATCCGACCT	791	
DB	463	CTGCCAGAGTGTGTGGGCTGTGTGGCCCTGCGGCCAGAGGTCTCTCCCTGGGATCCGACCT	522	
OY	792	CCCTTGGGCGCATCTCAAGAGCTGCCCCCTCTCACTACTCTGAGACTTTCAGAGTTCA	851	
DB	523	CCCTTGGGCGCATCTCAAGAGCTGCCCCCTCTCACTACTCTGAGACTTTCAGAGTTCA	582	
OY	852	CTGAGGGGCGCTGTCTCTCCCAAGTGTGTCCAGAGCTGTCCAGGCTGCGGCTCCCTCGACAGCTTC	911	
DB	583	CTGAGGGGCGCTGTCTCTCCCAAGTGTGTCCAGAGCTGTCCAGGCTGCGGCTCCCTCGACAGCTTC	642	
OY	912	TGGGACACCGGCGCCCTCTGCGCCACACCCCTCAACGCTCTTTCCTCAGACGCGCCCTGC	971	
DB	643	TGGGACACCGGCGCCCTCTGCGCCACACCCCTCAACGCTCTTTCCTCAGACGCGCCCTGC	702	
OY	972	CTCTAGAGAGCTGCTCGGGGCTGTTCACAGTGT-----TTTCCATCCACATAAATAGAGT--	1025	
DB	703	CTCTAGAGAGCTGCTCGGGGCTGTTCACAGTGTGGATATCATCATCCACATAAATAGAGTTA	762	
OY	1026	ATTCCCACTCTTATCTTACACTCCCGCCACGCGCCACTGTCCACTCA 1073		
DB	763	TTTCCCACTCTTATCTTACACTCCCGCCACGCGCCACTATCATCA 810		
<p>RESULT 8</p> <p>LOCUS B1870393 728 bp mRNA linear EST 11-OCT-2001</p> <p>LOCUS 60339564JF1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405459 5',</p>				

Accession	Version	KeyWords	Source	Organism
B070393	1	GI:16044066	mRNA sequence.	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
B070393	1	GI:16044066	EST.	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
B070393	1	GI:16044066	human.	Homo sapiens
B070393	1	GI:16044066	National Institutes of Health, Mammalian Gene Collection (MGC)	
B070393	1	GI:16044066	Unpublished (1999)	
B070393	1	GI:16044066	Contact: Robert Strausberg, Ph.D.	
B070393	1	GI:16044066	Email: cga@bbs-remail.nih.gov	
B070393	1	GI:16044066	Tissue Procurement: ATCC	
B070393	1	GI:16044066	CNA Library Preparation: Life Technologies, Inc.	
B070393	1	GI:16044066	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)	
B070393	1	GI:16044066	DNA Sequencing by: Incyte Genomics, Inc.	
B070393	1	GI:16044066	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:	
B070393	1	GI:16044066	http://image.llnl.gov	
B070393	1	GI:16044066	Plate: LAM12034 row: a column: 12	
B070393	1	GI:16044066	High quality sequence stop: 728.	
B070393	1	GI:16044066	Location/Qualifiers	
B070393	1	GI:16044066	1..728	
B070393	1	GI:16044066	/organism="Homo sapiens"	
B070393	1	GI:16044066	/db_xref="taxon:9606"	
B070393	1	GI:16044066	/clone="IMAGE:5405459"	
B070393	1	GI:16044066	/clone_lib="NIH-MGC_90"	
B070393	1	GI:16044066	/tissue_type="adenocarcinoma, cell_line"	
B070393	1	GI:16044066	/lab_host="DH10B (phage-resistant)"	
B070393	1	GI:16044066	/note="Organ: liver; Vector: pCMV-Sport6; Site_1: NotI; Site_2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC library."	
B070393	1	GI:16044066	BASE COUNT	
B070393	1	GI:16044066	125 a 240 c 227 g 136 t	
B070393	1	GI:16044066	ORIGIN	
B070393	1	GI:16044066	Query Match	
B070393	1	GI:16044066	Best Local Similarity 49.3%; Score 677.4; DB 13; Length 728;	
B070393	1	GI:16044066	Matches 724; Conservative 0; Pred. No. 1.6e-137;	
B070393	1	GI:16044066	Mismatches 1; Indels 6; Gaps 4	
B070393	1	GI:16044066	192 CCTGCGCTGGCCTTCCTCGGCCCTCTGCTGGCCCTGTCACTTTGGAGCGCGGCATC	
B070393	1	GI:16044066	1 CCTGCGCTGGCCTTCCTCGGCCCTCTGCTGGCCGTGTCACATTGGGGAGCGGGCATC	
B070393	1	GI:16044066	252 GCTGTCGCC---CAGGAGCTTGCCAGAGAGAGCTGTGGCAGAGAGACCAGACCC	
B070393	1	GI:16044066	61 GCTGTCGCCCCAGCAGGAGCCCTGCCAGAGAGAGCTGTGGCAGAGAGACCAGACCC	
B070393	1	GI:16044066	309 GTGCGAAGTGAATCCCGCAGACAGAAAGAACAGAGATCCTGCGCTTCTGAACGACT	
B070393	1	GI:16044066	121 GTGCGAAGTGAATCCCGCAGACAGAAAGAACAGAGATCCTGCGCTTCTGAACGACT	
B070393	1	GI:16044066	369 AGTTGGGCTCGCAGAGTCACTTAAGGCGGAGAAAACACGCGGCTCGAAGAGCGATCGC	
B070393	1	GI:16044066	181 AGTTGGGCTCGCAGAGTCACTTAAGGCGGAGAAAACACGCGGCTCGAAGAGCGATCGC	
B070393	1	GI:16044066	429 AGCCATTATGAAGTTTCATCCAGACCTTGACAGAGAGAGCGGAGAGAGGTGTGACCG	
B070393	1	GI:16044066	241 AGCCATTATGAAGTTTCATCCAGACCTTGACAGAGAGAGCGGAGAGAGGTGTGACCG	
B070393	1	GI:16044066	489 GACACTGAGTGGCTGGGAGAGAGCAGAAATCAACAGCTCACCGCTTCGCTACACCG	
B070393	1	GI:16044066	301 GACACTGAGTGGCTGGGAGAGAGCAGAAATCAACAGCTTCACCGCTTCGCTACACCG	
B070393	1	GI:16044066	549 CCAGATGCGGAGATTATAGTACCCGGGCTGGCTCTACTACTGTAAGTGTGAGGTGCA	
B070393	1	GI:16044066	361 CCAGATGCGGAGATTATAGTACCCGGGCTGGCTCTACTACTGTAAGTGTGAGGTGCA	
B070393	1	GI:16044066	609 CTTTATATAGGGGAGGCTGTCTACTGTAAGTGTGAGTGTGAGTGTGCTGCG	

Db	421	CCTTGTATAGAGGGGAAGGCGTGCTACTCCTAAGCTGACATTGCTGGTGCAGANTGGTGTGCTGCC	480
Oy	669	CCTGGCGTCCTGAGAAGAAATTTCTAGCCACTGCGGCCCACTTCCCTGGGGCCCCAGCTCG	728
Db	481	CCTGGCGTCCTGAGAAGAAATTTCTAGCCACTGCGGCCCACTTCCCTGGGGCCCCAGCTCG	540
Oy	729	CCTGGCGAGGTCGTGGGGCTTGGCCCTGGCGGCGAGGGTCCTCCCGCGGATCCGAC	788
Db	541	CCTGGCGAGGTCGTGGGGCTTGGCCCTGGCGGCGA-GGTCCTCCCGCGGATCCGAC	599
Oy	789	CTCCCTGGGGCCCATCTCAAGAGCTGCCCCCTCTCTCACACTTCCGAGACTTTCAGAGT	848
Db	600	CCTCCCTGGGGCCCATCTCAAGAGCTG-CCCCTCTCTCACACTTCCGAGACTTTCAGAGT	658
Oy	849	TCACTGAGGGGCGCTGTGCTCTCCCACTGCTGCCAGGCTGCGGCTCCCTCGACACT	908
Db	659	TCACTGAGGGGCGCTGTGCTCTCCCACTGCTGCCAGGCTGCGGCTCCCTCGACAGT	717
Oy	909	CTCTGGGCAACC	919
Db	718	CTCTGGGCAACC	728

RESULT 9
B1966060/c

LOCUS	B1966060	609 bp	mRNA	linear	EST 12-MAR-2002
DEFINITION	le72g04.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens				
CDNA CLONE IMAGE:	5672623 3' similar to TR:054907 054907 TNF-RELATED				
WEAK INDUCER OF APOPTOSIS :	mRNA sequence.				
ACCESSION	B1966060				
VERSION	B1966060.1	GI:16340465			
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
AUTHORS	Mammalia; Euthera; Primates; Carnivora; Hominiidae; Homo.				
	I (bases 1 to 609)				
	Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,				
	Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,				
	Hillier,L., Marra,M., Pape,D., Wylie,T., Martin,J., Blisstein,A.,				
	Schmitt,A., Theising,B., Ritter,E., Ronko,I., Bennett,J., Cardenas				
	,M., Gibbons,M., McCann,R., Cole,R., Tsagariswilli,R., Williams,T.,				
	Jackson,Y. and Bowers,Y.				
	Endocrine Pancreas Consortium				
JOURNAL COMMENT	Unpublished (2000)				
	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue				
	Endocrine Pancreas Consortium				
	Harvard University, Howard Hughes Medical Institute				
	Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,				
	MA 02138				
	Tel: 617-495-1812				
	Fax: 617-495-8557				
	Email: dmelton@biohp.harvard.edu				
	Library was constructed by Dr. Douglas Melton DNA sequencing by:				
	Washington University Genome Sequencing Center for information on				
	obtaining a clone please contact: Juliana Brown				
	(brownjefes.harvard.edu) This sequence now available from the IMAGE				
	consortium, for clone orders contact: info@image.llnl.gov				
	High quality sequence stop: 412.				
FEATURES					
source	Location/Qualifiers				
	1..609				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
	/clone="IMAGE:5672623"				
	/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"				
	/sex="Both"				
	/tissue_type="Islets of Langerhans"				
	/dev_stage="Adult"				
	/lab_host="DH10B"				
	/note="Organ: Pancreas; Vector: pSPORT1; Site:1; Not 1;				
	Site_2: Sal 1; Starting library constructed using				
	SuperScript Plasmid Library Kit (Life Technologies). cDNA				
	made by oligo-dT priming. Size-selected by column				

fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806. 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an Ecot of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

BASE COUNT 142 a 128 c 209 g 129 t 1 others
ORIGIN

Query Match 42.6%; Score 584.4; DB 13; Length 609;

Best Local Similarity 99.7%; Pred. No. 3.1e-117;

Matches 585; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 787 ACCCTCCCTGGGCCCCATCTCAAGGCTGCCCCCTCTCTACTACTCTGGAGCTCTCCAG 846
DB 609 ACCCTCCCTGGGCCCCATCTCAAGGCTGCCCCCTCTCTACTACTCTGGAGCTCTCCAG 550
QY 847 GTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGCTGCGGCGTCCCTCGACAG 906
DB 549 GTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGCTGCGGCGTCCCTCGACAG 490
QY 907 GTTCTGGGACCCGCTGCTGCTGCCCCAGCCCTGAGCCGCTTTGTCTCAGACCTGCC 966
DB 489 CTCTCTGGGACCCGCTGCTGCTGCCCCAGCCCTGAGCCGCTTTGTCTCAGACCTGCC 430
QY 967 CCTCCCTCTAGAGGCTGCTGCGGCGCTGCTTTCATCTCCACATTAATACAGTA 1026
DB 429 CTCTCCCTCTAGAGGCTGCTGCGGCGCTGCTTTCATCTCCACATTAATACAGTA 370
QY 1027 TTCCCACTCTTATCTTACAACTCCCCACAGCCCACTCTCCACATCTAGCTCCCAAT 1086
DB 369 TTCCCACTCTTATCTTACAACTCCCCACAGCCCACTCTCCACATCTAGCTCCCAAT 310
QY 1087 CCTGACCTCTTTGAGGCGCCCAAGTATCTGACTCCCTCGGCGCAGACCCCGAGGCG 1146
DB 309 CCTGACCTCTTTGAGGCGCCCAAGTATCTGACTCCCTCGGCGCAGACCCCGAGGCG 250
QY 1147 ATTTGTTTACTGTACTCTGTGGGCAAGGATGGTCCAGAAAGCCCACTTACAGGACTA 1206
DB 249 ATTTGTTTACTGTACTCTGTGGGCAAGGATGGTCCAGAAAGCCCACTTACAGGACTA 190
QY 1207 AGAGGGGCTGAGCTGCGCGGCAAGAGCCAAAGAGACTGGGCTAGGCGCAGAGTTCCCA 1266
DB 189 AGAGGGGCTGAGCTGCGCGGCAAGAGCCAAAGAGACTGGGCTAGGCGCAGAGTTCCCA 130
QY 1267 AATGTGAGGGGCGAGAAACAAGCAAGCTCTCCCTTGAGAAATTCCTGTGATTTTAA 1326
DB 129 AATGTGAGGGGCGAGAAACAAGCAAGCTCTCCCTTGAGAAATTCCTGTGATTTTAA 70
QY 1327 AACAGATATTATTATTATTTATTTGACAAATCTTATAAATG 1373
DB 69 AACAGATATTATTATTATTTATTTATTTGACAAATCTTATAAATG 23
```

RESULT 10
AF163779 1027 bp DNA linear GSS 29-AUG-2000
LOCUS AF163779 Human Homo sapiens genomic clone BAC750E14, DNA sequence.
ACCESSION AF163779
VERSION AF163779.1 GI:5726439
KEYWORDS GSS.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
JOURNAL Cousin, P., Billotte, J., Chabert, P. and Shaw, P. H.
Physical map of 17p13 and the genes adjacent to p53
Genomics 63 (1), 60-68 (2000)

MEDLINE

20130114

Contact: Shaw PH

Experimental Oncology

Institute of Pathology

Rue du Bugnon 25, Lausanne, VD 1011, Switzerland

subclone-AB2R Asc-BamHI PSL1180

Class: BAC subclone.

Location/Qualifiers

1..1027

/organism="Homo sapiens"

/db_xref="taxon:9606"

/map="17p"

/clone="BAC750E14"

/clone_11b="Human"

/note="Vector: pMO5 Blue"

BASE COUNT 207 a 317 c 282 g 208 t 13 others

ORIGIN

Query Match 42.3%; Score 581.4; DB 17; Length 1027;

Best Local Similarity 98.0%; Pred. No. 1.5e-116;

Matches 577; Conservative 8; Mismatches 4; Indels 0; Gaps 0;

```
QY 785 GCACCTCCCTGGGCGCCATCTCAAGGCTGCCCCCTCTCTACTACTCTGGAGCTCTCCAG 844
DB 1 GCACCTCCCTGGGCGCCATCTCAAGGCTGCCCCCTCTCTACTACTCTGGAGCTCTCCAG 60
QY 845 AGTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGCTGCGGCGTCCCTCGACAG 904
DB 61 AGTTCACTAGAGGGGCGCTGCTCCACAGTGTCCAGGCTGCGGCGTCCCTCGACAG 120
QY 905 AGCTCTGTGGGACCCGCTGCTGCTGCCCCAGCCCTGAGCCGCTTTGTCTCAGACCTG 964
DB 121 AGCTCTGTGGGACCCGCTGCTGCTGCCCCAGCCCTGAGCCGCTTTGTCTCAGACCTG 180
QY 965 CCCCTCCCTCTAGAGGCTGCTGCGGCGCTGCTTTCATCTCCACATTAATACAGTA 1024
DB 181 CCCCTCCCTCTAGAGGCTGCTGCGGCGCTGCTTTCATCTCCACATTAATACAGTA 240
QY 1025 TATTCACACTTATCTTACAACTCCCCACAGCCCACTCTCCACATCTAGCTCCCA 1084
DB 241 TATTCACACTTATCTTACAACTCCCCACAGCCCACTCTCCACATCTAGCTCCCA 300
QY 1085 ATCCCTGACCTCTTGAAGGCGCCCAAGTATCTGACTCCCTCGGCGCAGACCCCGAGG 1144
DB 301 ATCCCTGACCTCTTGAAGGCGCCCAAGTATCTGACTCCCTCGGCGCAGACCCCGAGG 360
QY 1145 GCATTGTGTTCACTGACTCTGTGGGCAAGATGGTCCAGAAAGCCCACTTACAGGAGC 1204
DB 361 GCATTGTGTTCACTGACTCTGTGGGCAAGATGGTCCAGAAAGCCCACTTACAGGAGC 420
QY 1205 TAAGAGGGGCTGAGCTGCGCGGCAAGAGCCAAAGAGACTGGGCTAGGCGCAGAGTTCC 1264
DB 421 TAAGAGGGGCTGAGCTGCGCGGCAAGAGCCAAAGAGACTGGGCTAGGCGCAGAGTTCC 480
QY 1265 CAAATGTGAGGGGCGAGAAACAAGCAAGCTCTCCCTTGAGAAATTCCTGTGATTTT 1324
DB 481 CAAATGTGAGGGGCGAGAAACAAGCAAGCTCTCCCTTGAGAAATTCCTGTGATTTT 540
QY 1325 AAAACAGATATTATTATTATTTATTTGACAAATCTTATAAATG 1373
DB 541 AAAACAGATATTATTATTATTTATTTATTTGACAAATCTTATAAATG 589
```

RESULT 11

BO884231

LOCUS BO884231 940 bp mRNA linear EST 16-AUG-2002

DEFINITION AGENCOURT_8682031 lupskl_sciatic_nerve Homo sapiens cDNA clone

IMAGE:6197488 5', mRNA sequence.

ACCESSION BO884231

VERSION BO884231.1 GI:22276239

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 940)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9apbs@mail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
 Plate: L1AM13607 row: 3 column: 17
 High quality sequence stop: 453.

FEATURES
 source
 Location/Qualifiers
 1..940
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6197488"
 /clone_lib="Lupski-sciatic-nerve"
 /sex="male"
 /tissue_type="sciatic nerve"
 /dev_stage="adult, 70 yr"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; CDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TCGACCCAGCGGCGC-3' and 5'-GACTAGTCTAGTACGAGCGCGCCGCTT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."

BASE COUNT 170 a 325 c 255 g 190 t
 ORIGIN

Query Match 40.9%; Score 562; DB 14; Length 940;
 Best Local Similarity 96.5%; Pred. No. 2.6e-112;
 Matches 660; Conservative 0; Mismatches 15; Indels 9; Gaps 8;

QY 301 CAGGACCGGTGGAACATATCCAGAGAGAAAGCCAGATCTGGCGCTTCTCG 360
 DB 1 CAGAGCCCGTGGAACTAATCCAGACAGAAAGCCAGATCTGGCGCTTCTCG 60
 QY 361 AACGACTAGTTCGGCTCGCAGAGAGTCCACTAAAGCCGGAACACAGGGCTCGAGA 420
 DB 61 AACGACTAGTTCGGCTCGCAGAGAGTCCACTAAAGCCGGAACACAGGGCTCGAGA 120
 QY 421 GCGATCGGAGCCATTAGAGTTCATCCAGACTCGACAGAGAGAGCGGCGCTGAGGT 480
 DB 121 GCGATCGGAGCCATTAGAGTTCATCCAGACTCGACAGAGAGAGCGGCGCTGAGGT 180
 QY 481 GTGACGGGACAGTGTGCTGGAGAGAGCCAGAAATCAACAGCTCAGCCCTCGCGC 540
 DB 181 GTGACGGGACAGTGTGCTGGAGAGAGCCAGAAATCAACAGCTCAGCCCTCGCGC 240
 QY 541 TACAACGGCCAGATCGGGAGTTTATAGTCAACCCGGGCTGAGCTTACTACTGTATGT 600
 DB 241 TACAACGGCCAGATCGGGAGTTTATAGTCAACCCGGGCTGAGCTTACTACTGTATGT 300
 QY 601 CAGGTGACCTTGTAGTAGGAGGAGGCTGTACCTGAAGCTGGAGCTTCTGTGATGTGT 660
 DB 301 CAGGTGACCTTGTAGTAGGAGGAGGCTGTACCTGAAGCTGGAGCTTCTGTGATGTGT 360
 QY 661 GTGCTGGCCCTGCGCTGCTGGAGAGATTCACAGCACTGCGGCGAGTTCCTCGGCGCC 720
 DB 361 GTGCTGGCCCTGCGCTGCTGGAGAGATTCACAGCACTGCGGCGAGTTCCTCGGCGCC 420
 QY 721 CAGCTCCGCCCTCTG-CCAGGTGTCTGGGCTGTGGCCCTCGCGGCA-GGGTCTCTCCCTG 778

DB 421 CAGCTCCGCCCTCTG-CCAGGTGTCTGGGCTGTGGCCCTCGCGGCA-GGGGCGCCCTCTG 480
 QY 779 GGATCCGACACCTCCCTGGGCGCCATCTCAGGCTG-CCCTCTCTCAGCTAC-TCGCG 836
 DB 481 GGATCCGACACCTCCCTGGGCGCCATCTCAGGCTG-CCCTCTCTCAGCTAC-TCGCG 540
 QY 837 ACTCTTCAGGTTCACTAGAGGGGCGCTGTCCACAGTGT-CCAGGCTGCGCGC- 894
 DB 541 ACTCTTCAGGTTCACTAGAGGGGCGCTGTCCACAGTGT-CCAGGCTGCGCGC 600
 QY 895 TCCCTCGACAGCTTCT-TGGGACCCGCTCCCTTGGCGGCGCC--TCAGCCGCTCTT 951
 DB 601 TCCCTCGACAGCTTCTGTGGGAAACCCGCTCCCTCTGCGCCACCCCTCAGCGGCTTCTT 660
 QY 952 TGCTCCAGACCTGCGCTTCTCTCT 975
 DB 661 TGCTCCAGACCTGCGCTTCTCTCT 684

RESULT 12
 BM971606/c 568 bp mRNA linear EST 21-MAR-2002
 LOCUS BM971606
 DEFINITION UI-CF-EC1-ab1-p-06-0-UI.s1 UI-CF-EC1 Homo sapiens cDNA clone
 ACCESSION UI-CF-EC1-ab1-p-06-0-UI.3', mRNA sequence.
 VERSION BM971606
 KEYWORDS BM971606.1 GI:19589193
 EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 568)
 AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
 TITLE Normalization and subtraction: two approaches to facilitate gene discovery
 JOURNAL Genome Res. 6 (9), 791-806 (1996)
 MEDLINE 97044477
 COMMENT Contact: McCray, PB
 McCray Lab
 University of Iowa
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
 Tel: 319 356 4866
 Fax: 319 356 7171

Email: paul.mccray@uiowa.edu
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone distribution: Researchers may obtain clones from Research Genetics (www.resgen.com)
 The following repetitive elements were found in this cDNA sequence: 1-82, >AT-rich/low-complexity (matched complement)
 Seq primer: M13 FORWARD
 POLYA=Yes.

FEATURES
 source

Location/Qualifiers
 1..568
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="UI-CF-EC1-ab1-p-06-0-UI"
 /clone_lib="UI-CF-EC1"
 /tissue_type="Lung"
 /dev_stage="Adult and Fetal"
 /lab_host="DH10B (Life Technologies) (T1 phage resistant)"
 /note="Organ: Lung; Vector: pT73-Pac (Pharmacia) with a modified polylinker; Site_1: EcoR I; Site_2: Not I; UI-CF-EC1 is a normalized cDNA library containing the following tissue(s): Normal lung from adult and from fetal day 64, day 87, week 19 and week 42. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I

Db	596	AGGGAACCAAGGAGACTTGAGCCCTTAGGCACAGCATTCGCCAAAGTGGACGGCAGAAGAAA	655
Oy	1286	AAGACAAGCTCTCTCCCTTGACAAATTCCTGTGGAATTTTAAA	1327
Db	656	CAGGAAAAGACCTCCCATGAGAGATGCGCTTGAGATATGACAA	697
<hr/>			
RESULT 14			
LOCUS	B1824443	531 bp	mRNA EST 04-OCT-2001
DEFINITION	60303869JF1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5179510 5', mRNA sequence.		
ACCESSION	B1824443		
VERSION	B1824443.1	GI:15935993	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. CDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: InCyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LLML1447 row: j column: 23 High quality sequence stop: 529. Location/Qualifiers 1..531 /organism="Homo sapiens" /db_xref="taxon:9606" /clone_image:5179510" /clone_id="NIH_MGC_115" /lab_host="DH10B" /note="Organ: pooled brain, lung, testis; Vector: pCMV-Sport6; Site:1: NotI; Site:2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH_MGC Library."		
BASE COUNT	72 a 223 c 119 g 117 t		
ORIGIN			
Query Match	37.6%; Score 516.8; DB 13; Length 531;		
Best Local Similarity	99.4%; Pred. No. 1.7e-102;		
Matches 529; Conservative 0; Mismatches 2; Indels 1; Gaps 1			
Oy	661	GTCGTGCCCCCTGCGCTGCGTGAAGAAATTCAGCCATGCGGCCAGTTCCCTCGGGCCC	720
Db	1	GTGCTGGCCCTCGCTGCTGCGTGAAGAAATTCAGCCATGCGGCCAGTTCCCTCGGGCCC	60
Oy	721	CAGCTCCGCGCTCGCCAGGTGTGTGGGGGTTGGCCCTGCGGCCAGGGTCCTCCCGGGG	780
Db	61	CAGCTCCGCGCTGTGCAGAGTGTGTGGGGTGTGGCCCTGCGGCCA-GGTCTTCCCTGGCG	119
Oy	781	ATCCGACACCCTCCCTGCGGCCCATCTCAAGAGGTGCCCCCTTCCTCACTACTTGGAGCTC	840
Db	120	ATTCGACACCCTCCCTGCGGCCCATCTCAAGAGGTGCCCCCTTCCTCACTACTTGGAGCTC	179
Oy	841	ATTCGAGTTCACTGAGGGGCGCTGGTCTCCCACTGCTCCAGCTGCGGCTGCGGCTCCCT	900

[illegible]

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2003, 07:57:26 ; Search time 382.56 Seconds

(Without alignments)
8082.384 Million cell updates/sec

Title: US-09-245-198a-3

Perfect score: 1373
Sequence: 1 atgcatgttagcttga.....gacaaatgtgataatg 1373

Scoring table:

IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SID52/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*

23: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SID52/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1373	100.0	1373 19	AAV18600 Homo sapiens tumou
2	1345.8	98.0	1364 24	ABK34881 Human cDNA encodin
3	1325	96.5	1421 20	AAAX56000 Human tumour necro
4	1320.2	96.2	1353 21	AAAX9717 Human PRO207 cDNA
5	1320.2	96.2	1353 24	ABK40255 cDNA encoding huma
6	1326.4	89.3	1236 19	AAV47613 TNF related endoth
7	1226.4	89.3	1236 22	AAD04350 Human TREPA (TNF r
8	833.4	60.7	1030 20	AAAX3424 Human TNRL3 DNA.
9	629.2	45.8	898 22	AA503964 Expression vector

10	614.6	44.8	1168	19	AAV18599	Mus musculus tumou
11	519.2	37.8	701	20	AAAX23425	Mouse TNRL3 DNA.
12	245.4	17.9	282	16	AAAT2190	Human gene signatu
13	193.4	14.1	195	24	ABK29540	Colon adenocarcino
14	70.4	5.1	114955	20	AAAX53491	Human adenosine A1
15	61	4.4	1337	20	AAAX17263	Human gene express
16	60.6	4.4	1000	21	ABAA02484	Human colon cancer
17	60	4.4	60	24	ABNA1049	Human spliced tran
18	60	4.4	60	24	ABN58591	Human spliced tran
19	60	4.4	60	24	ABN58593	Human spliced tran
20	60	4.4	60	24	ABN58848	Human spliced tran
21	60	4.4	60	24	ABN58849	Human spliced tran
22	59.2	4.3	10732	21	AAA10594	Gene encoding a su
23	58	4.2	1218	21	AAA02488	Oligonucleotide fo
24	54.4	4.0	600	24	ABO52496	Oligonucleotide fo
25	54.4	4.0	600	24	ABO52497	Oligonucleotide fo
26	54.4	4.0	1327	24	ABO68452	Listeria monocytog
27	54.4	4.0	1593	21	AAAO2504	Human colon cancer
28	53.2	3.9	53522	24	AAD30228	Human PKD1 gene.
29	53.2	3.9	53526	19	AAAT94101	Human PKD1 gene.
30	53.2	3.9	53577	17	AAAT8551	Human PKD1 gene.
31	53.2	3.9	53577	19	AAAT94108	Human PKD1 gene.
32	50	3.6	50	20	AAAX56002	Human PKD1 locus b
33	50	3.6	50	21	AAA49732	Human mature FLINT
34	50	3.6	50	24	ABK40292	Human mature FLINT
35	50	3.6	201	8	AAAT70195	Oligonucleotide pr
36	50	3.6	2188	20	AAZ77506	Streptomyces prote
37	49.6	3.6	1517	21	AAAT2028	Human ovarian tumo
38	49.2	3.6	825	20	AAZ25378	CDNA encoding huma
39	49.2	3.6	825	21	AAAS1078	Human mFLINT #2 nu
40	49.2	3.6	936	20	AAZ25376	Human mature FLINT
41	49.2	3.6	936	21	AAAB8731	Human FINT #2 nuc
42	49.2	3.6	936	21	AAA51076	Human FINT #2 nuc
43	49.2	3.6	936	22	AAAD07385	Human FINT coding
44	49.2	3.6	936	22	AAFB9920	Human FINT coding
45	49.2	3.6	936	22	AAFB4738	Nucleotide sequenc

ALIGNMENTS

RESULT 1	
AAV18600	AAV18600 standard; cDNA: 1373 BP.
ID	AAV18600;
AC	AAV18600;
XX	
DT	21-JUL-1998 (first entry)
XX	
DE	Homo sapiens tumour necrosis factor related ligand (TNFRL) gene.
XX	
KW	TNRL: tumour necrosis factor related ligand; tnfr: treatment;
KW	cancer; autoimmune disease; immune system; stimulation; suppression;
KW	graft rejection; ds..
XX	
OS	Homo sapiens.
XX	
FT	key
FT	1..852
FT	Location/Qualifiers
FT	/*tag= a
FT	/note= "tumour necrosis factor related ligand"
XX	
PN	W09805783-A1.
XX	
PD	12-FEB-1998.
XX	
PF	07-AUG-1997;
XX	97WO-US13945.
XX	
PR	18-MAR-1997;
XX	97US-0040820.
PR	07-AUG-1996;
XX	96US-0023541.
PR	18-OCT-1996;
XX	96US-0028515.
XX	
PA	(BIOD) BIOGEN INC.

PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.
 XX Browning JL, Chicheportiche Y;
 XX WPI, 1998-145619/13.
 DR P-PSDB; AAM47525.
 XX Tumour necrosis factor related ligand - useful for, e.g. treating
 PT cancer, auto-immune disease and immune responses to tissue grafts
 XX
 PS Claim 2; Pages 48-50; 69pp; English.
 XX
 CC The sequence is that encoding human tumour necrosis factor related
 CC ligand (TRELL). TRELL or active fragments can be included with a
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune
 CC diseases or immune responses to tissue grafts, or to stimulate or
 CC suppress the immune system. It is useful to screen for TRELL
 CC receptors, by labelling with a detectable label and screening
 CC compositions for binding. Agents interfering with TRELL-receptor
 CC binding can also be screened for, can then be administered,
 CC optionally with interferon- gamma, to induce cell death or
 CC treat, suppress or alter immune responses (especially involving human
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its
 CC receptor. The DNA sequence can be used in gene therapy for
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
 CC autoimmune and inflammatory diseases or inherited genetic disorders,
 CC by introducing into cells, and expressing, therapeutically effective
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
 CC It may also be of use in the preparation of prepare probes for
 CC screening natural/synthetic DNAs for TRELL-encoding sequences
 CC and for antisense therapy.
 CC
 XX
 SQ Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;
 Query Match 100.0%; Score 1373; DB 19; Length 1373;
 Best Local Similarity 100.0%; Pred. No. 6.9e-283;
 Matches 1373; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 1 ATGTCAATTGTAGACTTTGAATTTCCGCGCGGCTCCCTCCCGAGTCTTCGGG 60
 1 ATGTCAATTGTAGACTTTGAATTTCCGCGCGGCTCCCTCCCGAGTCTTCGGG 60
 61 TCCCGGATGGGGGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 61 TCCCGGATGGGGGGGGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 121 AGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
 121 AGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
 181 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 240
 181 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 240
 241 AGCCGGGATGCGTGTCCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 300
 241 AGCCGGGATGCGTGTCCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 241 AGCCGGGATGCGTGTCCGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 300
 301 CAGGACCCGTCGGAAGTATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 301 CAGGACCCGTCGGAAGTATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 361 AACCGACTAGTTCGGCTCGCAGAGAGTCACTAAAGGCGGAGAAACAGGCGCTCG 420
 361 AACCGACTAGTTCGGCTCGCAGAGAGTCACTAAAGGCGGAGAAACAGGCGCTCG 420
 421 GCGATGCGAGCCATTTATGAGTTTCATCCAGAGCTGGAGAGAGAGAGAGAGAG 480
 421 GCGATGCGAGCCATTTATGAGTTTCATCCAGAGCTGGAGAGAGAGAGAGAGAG 480
 481 GTGAGAGGAGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
 481 GTGAGAGGAGAGTGTGCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540

QY 541 TACACCGCAGATCGGGAGTATATAGTACACCGGGCTGGGCTACTACTACTGT 600
 DB 541 TACACCGCAGATCGGGAGTATATAGTACACCGGGCTGGGCTACTACTACTGT 600
 QY 601 CAGGTGACATTTGATAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
 DB 601 CAGGTGACATTTGATAGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
 QY 661 GTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 720
 DB 661 GTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 720
 QY 721 CAGCTCGGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
 DB 721 CAGCTCGGCTCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
 QY 781 ATCCGACCCCTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 DB 781 ATCCGACCCCTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 QY 841 TTCCAGGTTCACTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
 DB 841 TTCCAGGTTCACTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
 QY 901 CGACAGCTCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
 DB 901 CGACAGCTCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
 QY 961 CTTGCCCCCTCCCTCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
 DB 961 CTTGCCCCCTCCCTCTAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1020
 QY 1021 ACAGTATTCACCTTATCTATCTACATCCCGACGCGCTCTCTCTCTCTCTCT 1080
 DB 1021 ACAGTATTCACCTTATCTATCTACATCCCGACGCGCTCTCTCTCTCTCTCT 1080
 QY 1081 CCCAATCCCTGACCTCTTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
 DB 1081 CCCAATCCCTGACCTCTTGTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1140
 QY 1141 CAGGCACTGTCTTCACTGACTCTGTGGGCAAGAGAGGAGGAGGAGGAGGAGG 1200
 DB 1141 CAGGCACTGTCTTCACTGACTCTGTGGGCAAGAGAGGAGGAGGAGGAGGAGG 1200
 QY 1201 GCACCTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
 DB 1201 GCACCTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
 QY 1261 TTCCCAATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
 DB 1261 TTCCCAATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320
 QY 1321 TTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1373
 DB 1321 TTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1373

RESULT 2
 ABR34881
 ID ABR34881 standard; cDNA; 1364 BP.
 XX ABR34881;
 AC
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human cDNA encoding secreted protein #19.
 XX
 KW Human: secreted protein; gene; ss; nutritional supplement; haemophilia;
 KW viral infection; bacterial infection; fungal infection; diabetes; asthma;
 KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
 KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;

KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
 KW tissue regeneration; wound healing; burn; haematopoiesis;
 KW myeloid cell deficiency; lymphoid cell deficiency.
 OS Homo sapiens.
 XX
 XX
 PN WO200177288-A2.
 XX
 XX 18-OCT-2001.
 PD
 XX 29-MAR-2001; 2001WO-US10224.
 XX
 XX 06-APR-2000; 2000US-195582P.
 PR
 XX (GEMY) GENETICS INST INC.
 XX
 PA
 PI Wong GG, Clark HF, Fechtel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Gulkota K, Graham JR;
 PI XX
 DR WPI; 2002-179321/23.
 XX
 XX
 PT Five hundred and ninety two polynucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for
 PT treating immune deficiencies and disorders such as autoimmune disorders
 PT
 PS Claim 1; Page 82; 372pp; English.

CC The invention relates to 592 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins. The polynucleotides can be used as probes for the
 CC identification and isolation of full length cDNA and genomic DNA. The
 CC polynucleotides and proteins can also be used as nutritional supplements.
 CC The proteins are useful in the treatment of various immune deficiencies
 CC and disorders such as viral infections, bacterial infections, fungal
 CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions
 CC and conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilias),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment of
 CC burns, incisions and ulcers. The proteins are also useful for regulating
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
 CC Sequences ABK34863-ABK35354 represent polynucleotides of the invention.
 XX

Query Match	98.0%	Score 1345.8	DB 24	Length 1364
Best Local Similarity	99.9%	Pred. No. 4.3e-273		
Matches 1347	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

0Y	25	TCGGCCCGCCGGCTCCCCCTCCGCCGATACCTCCGGATCCCGGGATGGGGGGGGGGTAG	84
Db	5	TCGGCCCGCCGGCTCCCCCTCCGCCGATACCTCCGGATCCCGGGATGGGGGGGGGGTAG	64
0Y	85	CAGGCACAGCCCCCGCCCATGACCGCCCTTCGAGCCAGAGAGCGAGAGGGGCGCCGG	144
Db	65	CAGGCACAGCCCCCGCCCATGACCGCCCTTCGAGCCAGAGAGCGAGAGGGGCGCCGG	124
0Y	145	GGGGAGCCGGGACACCGCCCTCTGTCCCGCTCGCGATGGGGCCCGGGCCCTGGGCGCC	204
Db	125	GGGGAGCCGGGACACCGCCCTCTGTCCCGCTCGCGATGGGGCCCGGGCCCTGGGCGCC	184
0Y	205	TGCCTCGGCTCTGTGTGGCCGCTGTGATTTGGGGAGCCGGGATGAGTGTCCGCCAG	264
Db	185	TGCCTCGGCTCTGTGTGGCCGCTGTGATTTGGGGAGCCGGGATGAGTGTCCGCCAG	244
0Y	265	GAGCCTGGCCAGAGGAGACTGTTGGACAGAGAGACACGACCCCTGGGAATGAATCCC	324
Db	245	GAGCCTGGCCAGAGGAGACTGTTGGACAGAGAGACACGACCCCTGGGAATGAATCCC	304
0Y	325	CAGACGAGAAAGCCAGGATCTCGCGCTTTCTCGAACCGACTAGTTTCGGCCCTCCGAGA	384

Db	305	CAGACAGAAAGAAAGCCAGAGATCCTGGCCCTTTTCTTGACCAGACTAGTTGGCCCTGGCAGA	364
QY	385	AGTGCACCTTAAAGGCGGAAAAACACGGGCTTCGAAGAGCGATTCGACGCCCATTTATGAAGTT	444
Db	365	AGTGCACTTAAAGGCGGAAAAACACGGGCTTCGAAGAGCGATTCGACGCCCATTTATGAAGTT	424
QY	445	CATCCAGCACTTGCACAGAGACGAGCGCCAGGACAGGTGTGGACGGGACACTAGAGTGGCTGG	504
Db	425	CATCCACCACTTGCACAGAGACGAGCGCCAGGACAGGTGTGGACGGGACACTAGAGTGGCTGG	484
QY	505	GAGGAAGCCACAATCAACAGATCCAGCCCTCGCTGACAAACCGGCAGATTCGAGGGAGTTT	564
Db	485	GAGGAAGCCACAATCAACAGATCCAGCCCTCGCTGACAAACCGGCAGATTCGAGGGAGTTT	544
QY	555	ATAGTCACCCGGGCTGGGCTCTACTACTGTGTGACAGGTGCATTTGATGAGGGGAAG	624
Db	545	ATAGTCACCCGGGCTGGGCTCTACTACTGTGTGACAGGTGCATTTGATGAGGGGAAG	604
QY	625	GCTGTCTACTGAAGCTGAGACTTGTGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	684
Db	605	GCTGTCTACTGAAGCTGAGACTTGTGTGTGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	664
QY	685	GAATTCTCAGCCACTGCGGGCAGTTCCTCGGGGCCAGCTCCGGCTCTGCGCAGGAGTCT	744
Db	665	GAATTCTCAGCCACTGCGGGCAGTTCCTCGGGGCCAGCTCCGGCTCTGCGCAGGAGTCT	724
QY	745	GGGCTGTTGGCCCTGGGCGCCAGGGCTCCCTCGGATCGCACCCCTCCCTGGGCCCAT	804
Db	725	GGGCTGTTGGCCCTGGGCGCCAGGGCTCCCTCGGATCGCACCCCTCCCTGGGCCCAT	784
QY	805	CTCAAGGCTGGCCCCCTTCCTCACTACTTGTGGACTCTTCCAGGTCTACTGAGGGGCGCTG	864
Db	785	CTCAAGGCTGGCCCCCTTCCTCACTACTTGTGGACTCTTCCAGGTCTACTGAGGGGCGCTG	844
QY	865	GTCTCCCAAGTGTCACAGGCTGCGGGGTCCCGCTGCAGAGTCTCTGGGACCGCGTC	924
Db	845	GTCTCCCAAGTGTCACAGGCTGCGGGGTCCCGCTGCAGAGTCTCTCTGGGACCGCGTC	904
QY	925	CCCTCTGGCCCAACCCCTCAGCGGCTCTTGTGTCAGACCTGGCCCTCCCTCTAGAGAGTGC	984
Db	905	CCCTCTGGCCCAACCCCTCAGCGGCTCTTGTGTCAGACCTGGCCCTCCCTCTAGAGAGTGC	964
QY	985	CTGGGGCTGTTCACGCTGTTTTTCATCTCCACATATAATACAGTATTTCCCACTCTTACTTAC	1044
Db	965	CTGGGGCTGTTCACGCTGTTTTTCATCTCCACATATAATACAGTATTTCCCACTCTTACTTAC	1024
QY	1045	AACCTCCCCACCGGCCACTGCCACCTGCATAGCTGCCCAATCCCGAGCCCTTGTAGGGC	1104
Db	1025	AACCTCCCCACCGGCCACTGCCACCTGCATAGCTGCCCAATCCCGAGCCCTTGTAGGGC	1084
QY	1105	CCCACTGATCTCGACTCCGCCCTGGCCACAGACCCCGAGGGCATTTGTCTCACTGTACTC	1164
Db	1085	CCCACTGATCTCGACTCCGCCCTGGCCACAGACCCCGAGGGCATTTGTCTCACTGTACTC	1144
QY	1165	TGTGGGCAAGGATGAGTGTTCAGAGACCCCACTTCAAGGCACTAAGAGGGGCTGGACTGGC	1224
Db	1145	TGTGGGCAAGGATGAGTGTTCAGAGACCCCACTTCAAGGCACTAAGAGGGGCTGGACTGGC	1204
QY	1225	GCGAGGAAGCCAAAGAGACGTGGGCTTGCGCCAGAGAGTTCGCAAAATGTGAGGGGCGAGAA	1284
Db	1205	GCGAGGAAGCCAAAGAGACGTGGGCTTGCGCCAGAGAGTTCGCAAAATGTGAGGGGCGAGAA	1264
QY	1285	CAAGCAGAGCTCTCCCTTGAGAAATTCCTGTGGATTTTTTAAAAACAGATATTATTTTAT	1344
Db	1265	CAAGCAGAGCTCTCCCTTGAGAAATTCCTGTGGATTTTTTAAAAACAGATATTATTTTAT	1324
QY	1345	TATTTATTTGTGACAAAATGTTGATTAATGG 1373	
Db	1325	TATTTATTTGTGACAAAATGTTGATTAATGG 1353	

RESULT 3

[illegible]

Db	303	TGATCCCGACAGAAAGAAAGCCAGAGATCTGGCGCTTCTCGAACCAGACTAGTTGGCG	362
Qy	377	CTCGCAGAGTGCACCTAAAGCCCGGAAAAACAGGGCTCGAAGAGCATCCGAGCCATT	436
Db	363	CTCCACAAAGATGACACCTAAAGGCCCGGAAAAACAGGGCTCGAAGAGCATCCGAGCCATT	422
Qy	437	ATGAAGTTCATCCACGACCTGGACAGAGGAGAGCCGAGCAGTGTGGACGGGACAGTGA	496
Db	423	ATGAAGTTCATCCACGACCTGGACAGAGGAGAGCCGAGCAGTGTGGACGGGACAGTGA	482
Qy	497	GTCGCTGGGAGGAAGCCAGATCAACAGCTCCAGCCCTTCGCGTACAAACCGCAGATCG	556
Db	483	GTCGCTGGGAGGAAGCCAGATCAACAGCTCCAGCCCTTCGCGTACAAACCGCAGATCG	542
Qy	557	GGGAGTTATAGTACACCCGGGGCTGGGCTCTACTACCTGTACTGTCAAGGACCTTGATG	616
Db	543	GGGAGTTATAGTACACCCGGGGCTGGGCTCTACTACCTGTACTGTCAAGGACCTTGATG	602
Qy	617	AGGGGAAGGCTGTCTACCTGAACCTGGACTGTGCTGTGATGTGTGCTGGCCCTGCGCT	676
Db	603	AGGGGAAGGCTGTCTACCTGAACCTGGACTGTGCTGTGATGTGTGCTGGCCCTGCGCT	662
Qy	677	GCCCTGAGAGAAATTTCTACGCCACTGCGGCCAGTTCCTCGGGGCCCAAGCTCGGCTTGGC	736
Db	663	GCCCTGAGAGAAATTTCTACGCCACTGCGGCCAGTTCCTCGGGGCCCAAGCTCGGCTTGGC	722
Qy	737	AGGTGTCTGGGCTTTTGGCCCTCGGCCCAAGGTCCTCCCTGGGATCCGACACCTCCCT	796
Db	723	AGGTGTCTGGGCTTTTGGCCCTCGGCCCAAGGTCCTCCCTGGGATCCGACACCTCCCT	782
Qy	797	GGGCCCATCTCAGGCTGGCCCTTCTCAGCTACTTGGGACTCTTCCAGAGTTCACTAG	856
Db	783	GGGCCCATCTCAGGCTGGCCCTTCTCAGCTACTTGGGACTCTTCCAGAGTTCACTAG	842
Qy	857	GGGCCCTGAGTCTCCCAACAGTCTGCCAGGCTCCGGCTCCCTCGAGACAGCTCTTGGGC	916
Db	843	GGGCCCTGAGTCTCCCAACAGTCTGCCAGGCTCCGGCTCCCTCGAGACAGCTCTTGGGC	902
Qy	917	ACCCGGTCCCTCTGCCCCAACCTCAGCGGCTTTTGTCTCCAGAGCTCGCCCTCCCTCTA	976
Db	903	ACCCGGTCCCTCTGCCCCAACCTCAGCGGCTTTTGTCTCCAGAGCTCGCCCTCCCTCTA	962
Qy	977	GAGGCTGCTGGGCTGCTTCACGAGTTTTCATCCACATAAATACAGATATCCCACTCT	1036
Db	963	GAGGCTGCTGGGCTGCTTCACGAGTTTTCATCCACATAAATACAGATATCCCACTCT	1022
Qy	1037	TATCTTACAACCTCCCAACCGCCACTCTCCACCTCAGCTAGCTCCCAATCCCTGACCT	1096
Db	1023	TATCTTACAACCTCCCAACCGCCACTCTCCACCTCAGCTAGCTCCCAATCCCTGACCT	1082
Qy	1097	TTGAGGCCCCCAGTATCTGAGCTCCGCCCTGGGCGACAGACCCCAAGGGCATTTGTGTCA	1156
Db	1083	TTGAGGCCCCCAGTATCTGAGCTCCGCCCTGGGCGACAGACCCCAAGGGCATTTGTGTCA	1142
Qy	1157	CTGTACTCTGTGGGCAAGATGGGTCCAGAAAGCCCACTTAGCGCACTAAGAGGGGCTG	1216
Db	1143	CTGTACTCTGTGGGCAAGATGGGTCCAGAAAGCCCACTTAGCGCACTAAGAGGGGCTG	1202
Qy	1217	GACCTGGCGCAGGAGACCAAAAGAGCTGGGCTAGGCCAGGAGTTCCCAATGTGAGGG	1276
Db	1203	GACCTGGCGCAGGAGACCAAAAGAGCTGGGCTAGGCCAGGAGTTCCCAATGTGAGGG	1262
Qy	1277	GCGAGAAACAAGACAACCTCTCCCTTTGAGAAATTCCTCTGGATTTTAAACACATATTT	1336
Db	1263	GCGAGAAACAAGACAACCTCTCCCTTTGAGAAATTCCTCTGGATTTTAAACACATATTT	1322
Qy	1337	ATTTTATTTATTTATGTGACAAAATGTTGATTAATGG	1373
Db	1323	ATTTTATTTATTTATTTGACAAAATGTTGATTAATGG	1359

RESULT 4

AAA49717
 ID AAA49717 standard; cDNA; 1353 BP.
 AC AAA49717;
 XX
 XX 25-SEP-2000 (first entry)
 XX
 XX Human PRO207 cDNA clone DNA30879-1152.
 XX
 XX PRO207; human; antitumour; tumour; therapy; cytostatic;
 KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
 KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
 KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 58..807
 FT sig_peptide /*tag= a
 FT 58..177
 FT mat_peptide /*tag= b
 FT 178..804
 FT /*tag= c
 XX
 PN MO200037638-A2.
 XX
 PD 29-JUN-2000.
 XX
 PF 02-DEC-1999; 99MO-US28565.
 XX
 XX 22-DEC-1998; 98US-0113296.
 PR 08-MAR-1999; 99MO-US05028.
 PR 21-APR-1999; 99US-0130232.
 PR 28-APR-1999; 99US-0131445.
 PR 14-MAY-1999; 99US-0134287.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 15-SEP-1999; 99MO-US21090.
 PR 15-SEP-1999; 99MO-US21547.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
 PI Napier MA, Pitti RM, Wood WI;
 XX
 DR WPI: 2000-442668/38.
 DR P-PSDB; AAY95338.
 XX
 XX Novel composition to inhibit neoplastic cell growth or for treating
 PT tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
 PT PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO509 or
 PT PRO866
 PS
 PS Claim 20; Fig 3; 172pp; English.
 XX
 XX The present sequence is that of cDNA clone DNA30879-1152
 CC (ATCC 20335) encoding human PRO207 (see AAY95338), which shows
 CC homology to several members of the tumour necrosis factor family,
 CC especially human lymphotoksin (23.48). The cDNA was identified in a
 CC foetal kidney cDNA library following identification of an expressed
 CC sequence tag with homology to human Apo-2 ligand. A claimed method
 CC for inhibiting the growth of a tumour cell comprises exposing the
 CC tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,
 CC PRO328, PRO301, PRO526, PRO362, PRO509 or PRO866 (see
 CC AAY95337-49), their agonists or chimeric polypeptides incorporating
 CC them. The tumour is especially a cancer selected from breast,
 CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
 CC central nervous system cancer, melanoma and leukaemia. Nucleic
 CC acids encoding PRO179 etc. are used in the recombinant production
 CC of the antitumour polypeptides.
 CC
 XX Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;

Query Match 96.2%; Score 1320.2; DB 21; Length 1353;
 Best Local Similarity 99.8%; Pred. No. 1.2e-271;
 Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 49 CGATCCCTCGGGTCCGGGATGGGGGGGGGTGAGCAGGACAGCCCGCCCATG 108
 DB 1 CGATCCCTCGGGTCCGGGATGGGGGGGGGTGAGCAGGACAGCCCGCCCATG 60
 QY 109 GCGGCGCGTGGAGAGGCGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 168
 DB 61 GCGGCGCGTGGAGAGGCGGAGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 120
 QY 169 GTCCCGCTCGGCGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCT 228
 DB 121 GTCCCGCTCGGCGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCT 180
 QY 229 GTCAATTTGGGGAGCCGGGCGATCGCTGTCCGCCAGAGCCCTGCCAGAGAGCTGTG 288
 DB 181 GTCAATTTGGGGAGCCGGGCGATCGCTGTCCGCCAGAGCCCTGCCAGAGAGCTGTG 240
 QY 289 GCAGAGAGAGACAGACAGACCCGTCGGAATCGAATCCCAAGCAAGAAAGCAGATCT 348
 DB 241 GCAGAGAGAGACAGACAGACCCGTCGGAATCGAATCCCAAGCAAGAAAGCAGATCT 300
 QY 349 GCGCCTTTCCTGAGACCGACTAGTTCGGGCTTCGAGAAAGTGACCTAAAGGCGGAAAAA 408
 DB 301 GCGCCTTTCCTGAGACCGACTAGTTCGGGCTTCGAGAAAGTGACCTAAAGGCGGAAAAA 360
 QY 409 CGGGCTCGAAGAGCGATCGACAGCCCATTTATGACATCCACGACCTGACAGAGAGGA 468
 DB 361 CGGGCTCGAAGAGCGATCGACAGCCCATTTATGACATCCACGACCTGACAGAGAGGA 420
 QY 469 GCGCAGCAGAGTGTGAGACGGGACAGTAGTGTGGGAGAGGCAAGCAATTAACAGCTCC 528
 DB 421 GCGCAGCAGAGTGTGAGACGGGACAGTAGTGTGGGAGAGGCAAGCAATTAACAGCTCC 480
 QY 529 AGCCCTTCGCGCTCAACCGCCAGATGGGGAGTTTATGACACCGGGCTGGGCTTAC 588
 DB 481 AGCCCTTCGCGCTCAACCGCCAGATGGGGAGTTTATGACACCGGGCTGGGCTTAC 540
 QY 589 TACCTGTAAGTGTGAGTGTGACCTTTGATGAGGGAGAGGCTGTCTAGAGCTGGACTTG 648
 DB 541 TACCTGTAAGTGTGAGTGTGACCTTTGATGAGGGAGAGGCTGTCTAGAGCTGGACTTG 600
 QY 649 CTGGTGTAGTGTGTGAGGCGGCGCTGCGCTGCTGAGGAGATTCAGACACTGGGCGGAGT 708
 DB 601 CTGGTGTAGTGTGTGAGGCGGCGCTGCGCTGCTGAGGAGATTCAGACACTGGGCGGAGT 660
 QY 709 TCCCTCGGGGCCAGCTCCGCTTCGCGAGGTCTGTGGGCTGTGGGCTTCGCGGAGG 768
 DB 661 TCCCTCGGGGCCAGCTCCGCTTCGCGAGGTCTGTGGGCTGTGGGCTTCGCGGAGG 720
 QY 769 TCTTCCTCGGGAGATCGGACCCCTCCCTGGGCGCATCTCAAGGCTGCCCTTCTTACC 828
 DB 721 TCTTCCTCGGGAGATCGGACCCCTCCCTGGGCGCATCTCAAGGCTGCCCTTCTTACC 780
 QY 829 TACTTCGGAGCTTTCAGGTTTCACTGAGAGGGGCGCTGGGCTCCCAAGATGTGCCAGGCT 888
 DB 781 TACTTCGGAGCTTTCAGGTTTCACTGAGAGGGGCGCTGGGCTCCCAAGATGTGCCAGGCT 840
 QY 889 GCGGCGTCCCTCGACAGCTCTGTGGGACCCGGTCCCTTCGCGGACCCCTCAGCGCT 948
 DB 841 GCGGCGTCCCTCGACAGCTCTGTGGGACCCGGTCCCTTCGCGGACCCCTCAGCGCT 900
 QY 949 CTTTGTCTCAGACTGCCCTTCCTCTAGAGAGCTGTGGGCTGTTCAGTGTTCCTTTC 1008
 DB 901 CTTTGTCTCAGACTGCCCTTCCTCTAGAGAGCTGTGGGCTGTTCAGTGTTCCTTTC 960
 QY 1009 TCCCAATTAATACAGTATTCACCTTATCTTACAACTCCGACAGCCGACCTTCCTTTC 1068
 DB 961 TCCCAATTAATACAGTATTCACCTTATCTTACAACTCCGACAGCCGACCTTCCTTTC 1020
 QY 1069 CTTCACTAGCTCCCAATTCCTTGACCTTGTGAGGCGCCCACTGATCTGACTCCCTCTG 1128


```

|||||
Db 721 TCTTCCCGGAGATCCGACCTCCCTCCGCGCCATCTCAAGCTGCCCTTCTCTCAC 780
OY 829 TACTTGAGACTCTTCAGGTTCTACTAGAGGGCCCTGTCTCCCAAGTGTGCCAGGCT 888
Db 781 TACTTCGGACTCTTCAGGTTCTACTAGAGGGCCCTGTCTCCCAAGTGTGCCAGGCT 840
OY 889 GCCGGCTCCCTCCGACGCTCTCTGCGGACCCGGTCCCTCCGCGCCACCCGCTACCCGT 948
Db 841 GCCGGCTCCCTCCGACGCTCTCTGCGGACCCGGTCCCTCCGCGCCACCCGCTACCCGT 900
OY 949 CTTCGCTCCAGACGCTCCCTCCCTCTAGAGGCTCCCTGGGCTTTGACGTGTTTCCA 1008
Db 901 CTTCGCTCCAGACGCTCCCTCCCTCTAGAGGCTCCCTGGGCTTTGACGTGTTTCCA 960
OY 1009 TCCCACTAATAATACAGTATTCCTCACTTATCTAACAACCTCCCAAGCCGCTCTCCA 1068
Db 961 TCCCACTAATAATACAGTATTCCTCACTTATCTAACAACCTCCCAAGCCGCTCTCCA 1020
OY 1069 CCTCACTAGCTCCCAATCCCTGACCCCTTTGAGGCGCCCAAGTATCTGACTCCCGCTG 1128
Db 1021 CCTCACTAGCTCCCAATCCCTGACCCCTTTGAGGCGCCCAAGTATCTGACTCCCGCTG 1080
OY 1129 GCCACAGACCCCGCAGGATGTGTCTACTCTGTGCGCAAGATGGGTCCAGAAG 1188
Db 1081 GCCACAGACCCCGCAGGATGTGTCTACTCTGTGCGCAAGATGGGTCCAGAAG 1140
OY 1189 ACCCACTTCAGGACTAAGAGGGCTGAGCTGCGCGCAGAGCCCAAGAGACTGGGC 1248
Db 1141 ACCCACTTCAGGACTAAGAGGGCTGAGCTGCGCGCAGAGCCCAAGAGACTGGGC 1200
OY 1249 CTAGCGCAGAGTTCCCAATGTAGGGGCGAANAACAAGACCTCTCCCTTGAGAA 1308
Db 1201 CTAGCGCAGAGTTCCCAATGTAGGGGCGAANAACAAGACCTCTCCCTTGAGAA 1260
OY 1309 TTCCCTGTGATTTTAAACAGATATATTTTATTTATTTGACAAAATGTTGATA 1368
Db 1261 TTCCCTGTGATTTTAAACAGATATATTTTATTTATTTGACAAAATGTTGATA 1320
OY 1369 AATGG 1373
Db 1321 AATGG 1325

RESULT 6
AAV47613
ID AAV47613 standard; cDNA: 1236 BP.
AC AAV47613;
XX
DF 27-OCT-1998 (first entry)
XX
DE TNF related endothelium proliferative agent gene.
XX
KW ss: TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 1..750
FT /tag= a
FT /product= "TREPA"
XX
PM W09835061-A2.
XX
PD 13-AUG-1998.
XX
PF 12-FEB-1998; 98MO-US02859.
PR 10-FEB-1998; 98US-0021706.
PR 12-FEB-1997; 97US-0798692.
XX

```

```

PA (ABDO ) ABBOTT LAB.
XX
PI WILEY SR:
XX
DR WPI: 1998-447255/38.
DR P-PSDB: AAM29745.
XX
PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
PT treatment of autoimmune disease, tumours and inflammation
XX
PS Claim 11; Page 123-4; 142pp; English.
XX
CC The TNF-related endothelium proliferative agent (TREPA), or its
CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
CC promote wound healing or tissue grafting, by promoting vascularisation,
CC also to induce apoptosis for treating cancer and eliminating autoreactive
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
CC TREPA peptides can also be used to target cytotoxic agents or for
CC affinity isolation of the corresponding receptor, the nucleic acid for
CC which can be used to transform tumour cells to render them more
CC responsive to TREPA and to screen for TREPA mimics.
CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat
CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
CC vascularisation), inflammation or a wide range of autoimmune conditions,
CC conditions involving abnormal stimulation of epithelial cells (e.g.
CC atherosclerosis), for birth control (inhibiting ovulation and placental
CC formation) or other angiogenic conditions (e.g. ulcers).
XX
SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;

Query Match          89.3%; Score 1226.4; DB 19; Length 1236;
Best Local Similarity 99.5%; Pred. No. 1.1e-251;
Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 106 ATGCGCGCCGTGCGAGCCAGAGGGGCGCGGGGAGACCGGCGCTG 165
Db 1 ATGCGCGCCGTGCGAGCCAGAGGGGCGCGGGGAGACCGGCGCACCCGCTG 60
OY 166 CTGGTCCCGCTCGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGCTGCTGCGGCG 225
Db 61 CTGGTCCCGCTCGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGCTGCTGCGGCG 120
OY 226 GTGTGTCAGTTTGGGAGCGCGGCATCGCTGTCGCCCAAGAGCTCCCGCAGAGAGCTG 285
Db 121 GTGTGTCAGTTTGGGAGCGCGGCATCGCTGTCGCCCAAGAGCTCCCGCAGAGAGCTG 180
OY 286 GTGCGCAGAGAGACCGAGACCCGTGCGAATGTAATCCCGACAGAGAGAAACCGAGAT 345
Db 181 GTGCGCAGAGAGACCGAGACCCGTGCGAATGTAATCCCGACAGAGAGAAACCGAGAT 240
OY 346 CCTGCGCCTTTCCTGTAACCGATAGTTCGCGCTCGCAGAGAGGCAACCTTAAGCGCGGAA 405
Db 241 CCTGCGCCTTTCCTGTAACCGATAGTTCGCGCTCGCAGAGAGGCAACCTTAAGCGCGGAA 300
OY 406 ACACGGGCTCGAAGAGCCATTCGACGCCATTTGAAGTTTCATCCAGACTGGACAGGAC 465
Db 301 ACACGGGCTCGAAGAGCCATTCGACGCCATTTGAAGTTTCATCCAGACTGGACAGGAC 360
OY 466 GGAGCGCAGGACAGGTGTGAGCGGACAGTGTGCTGGGAGAGAACCGAATCAACAGC 525
Db 361 GGAGCGCAGGACAGGTGTGAGCGGACAGTGTGCTGGGAGAGAACCGAATCAACAGC 420
OY 526 TCGAGCCCTTCGCTGACAAACCGCAGATCGGGGAGTTATATGTACCCGGGCTGGGCTC 585
Db 421 TCGAGCCCTTCGCTGACAAACCGCAGATCGGGGAGTTATATGTACCCGGGCTGGGCTC 480
OY 586 TACTACTGTACTGTACAGTGTGACTTTGATGAGGGGAAAGCTGTACTGTGAAGCTGAG 645
Db 481 TACTACTGTACTGTACAGTGTGACTTTGATGAGGGGAAAGCTGTACTGTGAAGCTGAG 540
OY 646 TTGCTGTGTGATGTGTGTGCTGCGCTGCTGCTGAGAGAAATCTCAGCCACTGCGGCG 705
Db 541 TTGCTGTGTGATGTGTGTGCTGCGCTGCTGCTGAGAGAAATCTCAGCCACTGCGGCG 600

```

OY	ACTTCCCTCGGGGCCACAGCTCCGCTCCAGCTGCTCGAGTGTCTGGCTGTTTGACCCCTGGGGCCA	765
Db	601 ACTTCCCTCGGGGCCACAGCTCCGCTCCAGCTGCTCGAGTGTCTGGCTGTTTGACCCCTGGGGCCA	660
OY	766 GGGTCTCTCCCTCGGGATCCGACACCTCCCTGGGCCATCTCAAGGCTGCCCTTCCTC	825
Db	661 GGGTCTCTCCCTCGGGATCCGACACCTCCCTGGGCCATCTCAAGGCTGCCCTTCCTC	720
OY	826 ACCTACTTCGGAGCTCTCCAGGTTCACTGAGGGGGCCCTGGTCTCCCAAGTGTGCCAG	885
Db	721 ACCTACTTCGGAGCTCTCCAGGTTCACTGAGGGGGCCCTGGTCTCCCAAGTGTGCCAG	780
OY	886 GCTGCGGGCTCCCTCGACAGCTCTCTGGGACCCGGTCCCTCTGGCCCACTCAGC	945
Db	781 GCTGCGGGCTCCCTCGACAGCTCTCTGGGACCCGGTCCCTCTGGCCCACTCAGC	840
OY	946 GCTCTTTTGTCCAGACCTGCCCCCTCTTAAGAGCTGCTGGGCTGTTCACGTGTTTT	1005
Db	841 GCTCTTTTGTCCAGACCTGCCCCCTCTTAAGAGCTGCTGGGCTGTTCACGTGTTTT	900
OY	1006 CCATCCCAATAAATACAGTATTTCACACTTATCTTATCTTACACGTCGCCACCGCCACTCT	1065
Db	901 CCATCCCAATAAATACAGTATTTCACACTTATCTTATCTTACACACCCACCGCCACTCT	960
OY	1066 CCACCTCACTAGCTCCCCCAATCCCTGACCCCTTTGAGGGCCCAAGTATCTGACTCCCC	1125
Db	961 CCACCTCACTAGCTCCCCCAATCCCTGACCCCTTTGAGGGCCCAAGTATCTGACTCCCC	1020
OY	1126 CTGGGCACAGACCCCAAGGCAATTGTGTCTACTGTACTGTGAGGCAAGATGGTCCAG	1185
Db	1021 CTGGGCACAGACCCCAAGGCAATTGTGTCTACTGTACTGTGAGGCAAGATGGTCCAG	1080
OY	1186 AAGACCCCACTTCAGGCACTAAGAGGGGCTGAGCCTGGGCGCAGAGCAAGCAAGAGACTG	1245
Db	1081 AAGACCCCACTTCAGGCACTAAGAGGGGCTGAGCCTGGGCGCAGAGCAAGCAAGAGACTG	1140
OY	1246 GGGCTTAGGCGAGAGATGTCCCAATATGTGAGGGGCGAGAAACAAGACAGCTCTCCCTTGA	1305
Db	1141 GGGCTTAGGCGAGAGATGTCCCAATATGTGAGGGGCGAGAAACAAGACAGCTCTCCCTTGA	1200
OY	1306 GAATTCCTGTGTGATTTTAAACAAGATTTATTTT 1341	
Db	1201 GAATTCCTGTGTGATTTTAAACAAGATTTATTTT 1236	
RESULT 7		
AAD04350		
XX	AAD04350 standard; cDNA: 1236 BP.	
XX	AAD04350;	
AC		
XX	04-JUL-2001 (first entry)	
DT		
XX		
XX	Human TREPA (TNF related endothelium proliferative agent) cDNA.	
DE		
XX	Human: tumour necrosis factor: TNF; angiogenesis; wound healing;	
KW	TREPA; TNF related endothelium proliferative agent; tumour; metastasis;	
KM	grafting; vulnerability; ss.	
KW		
XX	Homo sapiens.	
OS		
XX		
XX	Key Location/Qualifiers	
FF	CDS 1..750	
FT	/*tag= a	
FT	/product= "Human TREPA (TNF related endothelium	
FT	proliferative agent)"	
XX	US6207642-B1.	
XX	27-MAR-2001.	
PD		
XX	26-JUN-1998; 98US-0105343.	
PF		

[illegible]

```

QY 706 AGTTCCTCGGGCCCACTCCGCTCTGCGAGGTGTGGGCTGTTGGCCCTGGGAGCA 765
    |||||||
Db 601 AGTTCCTCGGGCCCACTCCGCTCTGCGAGGTGTGGGCTGTTGGCCCTGGGAGCA 660
QY 766 GGGTCTCTCGGGGATCCGACCCCTCCCTGGGCCCCATCTCAAGGCTGCCCCCTTCTC 825
    |||||||
Db 661 GGGTCTCTCGGGGATCCGACCCCTCCCTGGGCCCCATCTCAAGGCTGCCCCCTTCTC 720
QY 826 ACCACTTCGGGAGCTTCCAGGTTCACGTAGAGGGGGCTGGGTCTCCCACTCCGCTCCAG 885
    |||||||
Db 721 ACCACTTCGGGAGCTTCCAGGTTCACGTAGAGGGGGCTGGGTCTCCCACTCCGCTCCAG 780
QY 886 GCTGCGGGCTCCCTCGACAGCTCTGTGGGACCCGGTCCCTCTGCCCCACCTCAGCC 945
    |||||||
Db 781 GCTGCGGGCTCCCTCGACAGCTCTGTGGGACCCGGTCCCTCTGCCCCACCTCAGCC 840
QY 946 GCTCTTTGCTCCAGACCTGCCCCCTCTGTAGAGGCTGCTGGGCTGTTCAAGTGT 1005
    |||||||
Db 841 GCTCTTTGCTCCAGACCTGCCCCCTCTGTAGAGGCTGCTGGGCTGTTCAAGTGT 900
QY 1006 CCATCCCAATAAATAGATGATTCCTCCACTTATCTTCAACTCCGCCACGCCCACTCT 1065
    |||||||
Db 901 CCATCCCAATAAATAGATGATTCCTCCACTTATCTTCAACAACCCACGCCCACTCT 960
QY 1066 CCACCTCACTAGCTCCCAATCCCTGACCTTTTGAGGCCGCCAGTGTGACTCCGCC 1125
    |||||||
Db 961 CCACCTCACTAGCTCCCAATCCCTGACCTTTTGAGGCCGCCAGTGTGACTCCGCC 1020
QY 1126 CTGGCCACAGACCCCGAGGCAATGTGTCTCACTGTCTGTGGGCAAGATGGTCCAG 1185
    |||||||
Db 1021 CTGGCCACAGACCCCGAGGCAATGTGTCTCACTGTCTGTGGGCAAGATGGTCCAG 1080
QY 1186 AAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGAGGAAAGCAAGAGACTG 1245
    |||||||
Db 1081 AAGACCCCACTTCAGGCACTAAGAGGGGCTGGACCTGGCGGAGGAAAGCAAGAGACTG 1140
QY 1246 GGCCTAGGCCAGAGGATTCCTCAATGTGAGGGCGGAGAAACAAGACACTCTCCCTTGA 1305
    |||||||
Db 1141 GGCCTAGGCCAGAGGATTCCTCAATGTGAGGGCGGAGAAACAAGACACTCTCCCTTGA 1200
QY 1306 GAATTCCTCTGTGATTTTAAACAGATATATTTT 1341
    |||||||
Db 1201 GAATTCCTCTGTGATTTTAAACAGATATATTTT 1236

RESULT 8
AAK23424
ID AAK23424 standard; DNA; 1030 BP.
XX
AC AAK23424;
XX
DT 18-JUN-1999 (first entry)
XX
DE Human TNRL3 DNA.
XX
KW Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;
    developmental abnormality; gestational abnormality; prostate cancer;
    AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
    cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
    apoptosis; human; ss.
XX
KM
XX
OS Homo sapiens.
XX
FH Key 1..627 Location/Qualifiers
FT CDS /*tag= a
FT /product= "TNRL3"
XX
PN MO9911791-A2.
XX
PD 11-MAR-1999.
XX
PF 04-SEP-1998; 98WO-US18393.

```

```

XX
PR 05-SEP-1997; 97US-0924634.
PA (UNIV ) UNIV WASHINGTON.
XX
PI Chaudhary PM;
XX
DR WPI: 1999-205191/17.
DR P-PSDB: AAM93590.
XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
PS Example VII; Fig 13A; 156pp; English.
XX
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: AP04, AP06, AP08 and AP09 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. AP04 is useful for diagnosing prostate cancer
CC by determining levels of AP04 in an individual. Prostate cancer can also
CC be treated using AP04 selective binding agents linked to a therapeutic
CC moiety. AP04 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. AP04 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in AP04
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using AP04 polypeptides/active
CC fragments and AP04 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of AP04 and detecting a change in level of AP04
CC activity. The method is performed in vivo or in vitro. AP0 polypeptides
CC are all useful as immunogens for preparing antibodies. AP04 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. AP08 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;

Query Match 60.7%; Score 833.4; DB 20; Length 1030;
Best Local Similarity 99.9%; Pred. No. 4.8e-168;
Matches 834; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 229 GTCAATTTGGGAGCGGGGCAATCGCTGTCGGCCAGAGAGCCGCCAGAGAGCTG 288
    |||||||
Db 1 GTCAATTTGGGAGCGGGGCAATCGCTGTCGGCCAGAGAGCTG 60
QY 289 GCAGAGAGAGACAGAGACCCGTCGAGACTGAATCCCAAGACAGAAAGCCAGAGATCT 348
    |||||||
Db 61 GCAGAGAGAGACAGAGACCCGTCGAGACTGAATCCCAAGACAGAAAGCCAGAGATCT 120
QY 349 GCGCTTTCTGAAACGACTAGTTGGCTCGCAGAACTGCACCTTAAAGCCGGA 408
    |||||||
Db 121 GCGCTTTCTGAAACGACTAGTTGGCTCGCAGAACTGCACCTTAAAGCCGGA 180
QY 409 CGGGCTGAAGAGGATGAGAGCCCATATGAAGTTCATCCACGACGCTGGAGAGAGGA 468
    |||||||
Db 181 CGGGCTGAAGAGGATGAGAGCCCATATGAAGTTCATCCACGAGCTGGAGAGAGAGGA 240
QY 469 GCGCAGCAGAGGTGTGAGCGGAGACAGTAGTGCTGGAGAGAAAGCCAGATCAACAGCTCC 528
    |||||||
Db 241 GCGCAGCAGAGGTGTGAGCGGAGACAGTAGTGCTGGAGAGAAAGCCAGATCAACAGCTCC 300
QY 529 AGCCCTGTGCGTACACCGCCAGATCGGGAGATTATAGTACCCGGGGCTGAGCTTAC 588
    |||||||
Db 301 AGCCCTGTGCGTACACCGCCAGATCGGGAGATTATAGTACCCGGGGCTGAGCTTAC 360
QY 589 TACCTGTACTGTACAGTCACTTTGATAGAGGAGAGGCTGTCTACTCAAGCTGACCTG 648
    |||||||
Db 361 TACCTGTACTGTACAGTCACTTTGATAGAGGAGAGGCTGTCTACTCAAGCTGACCTG 420
QY 649 CTGCTGATGATGCTGCGCTGCGCTGCGTGCAGAGAAATTCAGCCACTGCGGCAGT 708

```

```

DB 421 CTGCTGAGATGTCGTGCTGCGCTGCGCTGAGAGAAATTCACAGCCACTCGGCGCAGT 480
OY 709 TCCCTGCGGCCCCAGCTCGGCTCGGCCAGGTCTGCGGCTGTGGCCCTCGGCGCCAGGG 768
DB 481 TCCCTGCGGCCCCAGCTCGGCTCGGCCAGGTCTGCGGCTGTGGCCCTCGGCGCCAGGG 540
OY 769 TCCCTGCGGCCCCAGCTCGGCTCGGCCAGGTCTGCGGCTGTGGCCCTCGGCGCCAGGG 828
DB 541 TCCCTGCGGCCCCAGCTCGGCTCGGCCAGGTCTGCGGCTGTGGCCCTCGGCGCCAGGG 600
OY 829 TACTGCGACTCTTCCAGCTTCACTGAGGCGCCCTGCTCTCCACAGTGTCCAGGCT 888
DB 601 TACTGCGACTCTTCCAGCTTCACTGAGGCGCCCTGCTCTCCACAGTGTCCAGGCT 660
OY 889 GCGGCGTCCCTCGACAGCTCTGCGGCGCCGCTGCTGCGGCGCCCTCGGCGCCAGGCT 948
DB 661 GCGGCGTCCCTCGACAGCTCTGCGGCGCCGCTGCTGCGGCGCCCTCGGCGCCAGGCT 720
OY 949 CTTTGGCTCCAGACCTGCGCCCTGCTCTAGAGGCTGCTGCGGCTGTTCACGTCTTTTCA 1008
DB 721 CTTTGGCTCCAGACCTGCGCCCTGCTCTAGAGGCTGCTGCGGCTGTTCACGTCTTTTCA 780
OY 1009 TCCCATATTAATACATATTTCCACCTTATCTTACAACTCCCGCCAGCCGCTACT 1063
DB 781 TCCCATATTAATACATATTTCCACCTTATCTTACAACTCCCGCCAGCCGCTACT 835

RESULT 9
AAS03964
ID AAS03964 standard; DNA; 898 BP.
XX
AC AAS03964;
XX
DT 26-SEP-2001 (first entry)
XX
DE Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA.
XX
KW TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;
KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
KW rubrosis; uveitis; macular degeneration; arthritis; rheumatism; ds;
KW corneal graft neovascularisation; psoriasis; benign tumour; haemophilic joint;
KW malignant tumour; sarcoma; carcinoma; benign tumour; metastatic condition;
KW preneoplastic condition; myocardial angiogenesis; wound granulation;
KW scleroderma; vascular adhesion; telangiectasia; ischaemia; human;
KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;
KW peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAKR;
KW fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 52..873
FT /tag= "Fusion protein comprising a growth hormone
FT /product= leader, a leucine zipper multimerisation
FT domain, and human TWEAK extracellular
FT domain"
XX
PN WO200145730-A2.
XX
PD 28-JUN-2001.
XX
PF 19-DEC-2000; 2000WO-US34755.
XX
PR 20-DEC-1999; 99US-0172878.
XX 10-MAY-2000; 2000US-0203347.
XX (IMMUNEX CORP.
XX
PI WILEY SR;

```

```

XX WPI: 2001-417975/44.
DR P-PSDB: AAU03499.
XX
PT Modulating angiogenesis in a mammal for treating diseases mediated by
PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
PT peripheral tissue, by administering antagonist or agonist of TWEAK
PT receptor
XX
PS Example 1; Page 39-40; 46pp: English.
XX
CC The sequence represents a DNA from the expression vector
CC pDC409-LZ-TWEAK, which encodes a fusion protein comprising a growth
CC hormone leader, a leucine zipper multimerisation domain, and the
CC extracellular domain of human TWEAK. The fusion protein was used in
CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones
CC from a COS cell human cDNA library. The TWEAK protein is a
CC member of the tumour necrosis factor (TNF) family and induces
CC angiogenesis. TWEAKR may therefore be used to screen for and develop
CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be
CC used in the treatment and diagnosis of human disease. The disorders
CC mediated by angiogenesis include ocular disorders characterised by ocular
CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,
CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
CC rubrosis, uveitis, macular degeneration and corneal graft
CC neovascularisation, and inflammatory diseases such as arthritis,
CC rheumatism and psoriasis. Other treatable diseases include malignant and
CC metastatic conditions such as sarcomas and carcinomas, benign tumours and
CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
CC scleroderma, vascular adhesions, atherosclerotic plaque
CC neovascularisation, telangiectasia, wound granulation, coronary
CC atherosclerosis, peripheral atherosclerosis and ischaemia.
CC
SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other;
XX
Query Match 45.8%; Score 629.2; DB 22; Length 898;
Best Local Similarity 99.5%; Pred. No. 1,4e-124;
Matches 631; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
OY 232 AGTTTGGGAGCGGGGATCGCTGTCGCCGCGAGAGCTGCCAGAGAGAGTGTGCGCA 291
DB 250 AGTTTGGGAGCGGGGATCGCTGTCGCCGCGAGAGCTGCCAGAGAGAGTGTGCGCA 309
OY 292 GAGGAGACCGAGACCGCTGCGAAGTATCCCGACAGAAAGAACCGAGATCTGCG 351
DB 310 GAGGAGACCGAGACCGCTGCGAAGTATCCCGACAGAAAGAACCGAGATCTGCG 369
OY 352 CCTTCTGACCGAGTACTGTCGGCTCGCAGAGAGCGACTTAAGGCGGAAACACGG 411
DB 370 CCTTCTGACCGAGTACTGTCGGCTCGCAGAGAGCGACTTAAGGCGGAAACACGG 429
OY 412 GCTGAAGAGCGATCGACGCCCATTTATGAAGTTTCATCCAGACTGAGACGAGGCG 471
DB 430 GCTGAAGAGCGATCGACGCCCATTTATGAAGTTTCATCCAGACTGAGACGAGGCG 489
OY 472 CAGGAGGTGTGAGCGGAGCAGTGTGCTGGGAGAGCCAGATCAACAGCTCCAGC 531
DB 490 CAGGAGGTGTGAGCGGAGCAGTGTGCTGGGAGAGCCAGATCAACAGCTCCAGC 549
OY 532 CCTGCGCTTACAAACCGGAGATCGGGAGTTTATGTCACCGGCGCTGCTTACTAC 591
DB 550 CCTGCGCTTACAAACCGGAGATCGGGAGTTTATGTCACCGGCGCTGCTTACTAC 609
OY 592 CTGACTGTCAAGTGCATTTGATGAGGGAAGGCTGTACTGAAAGCTGAGCTTGTG 651
DB 610 CTGACTGTCAAGTGCATTTGATGAGGGAAGGCTGTACTGAAAGCTGAGCTTGTG 669
OY 652 GTGAGATGTGTGCTGCGCTGCGCTGCGAGAGATTTTCAGCCACTCGGCGCACTTC 711
DB 670 GTGAGATGTGTGCTGCGCTGCGCTGCGAGAGATTTTCAGCCACTCGGCGCACTTC 729
OY 712 CTGCGGCGGCGAGCTCGGCTGCGAGGTGTGCGGCTGTGGCCCTGCGGCGAGGCTCC 771

```

Db	730	CTGGGGCCCCAGCTCCGCTTCGTGGCAGTGTGTGGGCTGTGGGCCCTGCGGCCACAGGTGC	789
Oy	772	TCCTCTGGGATCCGACCCCTCCCTGGAGCCATCTCAAGGCTGCCCCCTTCTCAGCTAC	831
Db	790	TCCTCTGGGATCCGACCCCTCCCTGGAGCCATCTCAAGGCTGCCCCCTTCTCAGCTAC	849
Oy	832	TTCCGACTCTTCACAGTTCTACTGAGGGGCCCTGG	865
Db	850	TTCCGACTCTTCACAGTTCTACTGAGGGGCCCTGG	883
RESULT 10			
AAV18599			
ID	AAV18599	standard; cDNA: 1168 BP.	
AC	AAV18599;		
XX			
DT	21-JUL-1998	(first entry)	
XX			
De	Mus musculus	tumour necrosis factor related ligand (TREL) gene.	
XX			
KW	TREL; tumour necrosis factor related ligand; tnf; treatment;		
KW	cancer; autoimmune disease; immune system; stimulation; suppression;		
KW	graft rejection; ds.		
XX			
OS	Mus musculus.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	2..679	
FT		/*tag= a	
FT		/note= "tumour necrosis factor related ligand"	
XX			
PN	W09805783-A1.		
XX			
PD	12-FEB-1998.		
XX			
PE	07-AUG-1997;	97WO-US13945.	
XX			
PR	18-MAR-1997;	97US-0040820.	
PR	07-AUG-1996;	96US-0023541.	
PR	18-OCT-1996;	96US-0028515.	
XX			
PA	(BIOJ) BIOGEN INC.		
PA	(UYGE-) UNIV GENEVA FACULTY MEDICINE.		
XX			
PI	Browning JL, Chicheportliche Y;		
XX			
DR	WPI: 1998-145619/13.		
DR	P-PSDB; AAW47524.		
XX			
PT	Tumour necrosis factor related ligand - useful for, e.g. treating		
PT	cancer, auto-immune disease and immune responses to tissue grafts		
XX			
PS	Claim 2; Pages 45-46; 69pp; English.		
XX			
CC	The sequence is that encoding mouse tumour necrosis factor related		
CC	ligand (TREL). TREL or active fragments can be included with a		
CC	carrier in pharmaceutical compositions to treat cancer, autoimmune		
CC	diseases or immune responses to tissue grafts, or to stimulate or		
CC	suppress the immune system. It is useful to screen for TREL		
CC	receptors, by labelling with a detectable label and screening		
CC	compositions for binding. Agents interfering with TREL-receptor		
CC	binding can also be screened for, can then be administered,		
CC	optionally with interferon- gamma, to induce cell death or		
CC	treat, suppress or alter immune responses (especially involving human		
CC	adenocarcinoma cells) involving a signal pathway between TREL and its		
CC	receptor. The DNA sequence can be used in gene therapy for		
CC	TREL-related disorders in mammals (especially humans), e.g. tumours,		
CC	autoimmune and inflammatory diseases or inherited genetic disorders,		
CC	by introducing into cells, and expressing, therapeutically effective		
CC	amounts of a vector, e.g. a virus comprising a gene encoding TREL.		
CC	It may also be of use in the preparation of prepare probes for		
CC	screening natural/synthetic DNAs for TREL-encoding sequences		

CC	and for antisense therapy.
XX	
SQ	Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;
	Query Match 44.8%; Score 614.6; DB 19; Length 1168; Best Local Similarity 75.8%; Pred. No. 1,8e-121; Matches 906; Conservative 0; Mismatches 219; Indels 70; Gaps 9.
QY	178 GCGCTGGGCTGGGCTTGCGCTGGCCTTCCTCGGCCTCTCTGCGCCGTGGTCAAGTTG 237 DB 2 GTGCTAGAGCTGGGCTGGCGCTGGCGCTGCCCTTGGGCTCTCTGCTCTGCTGGTGGTAGACCG 61
OY	238 GGGAGCGGGCATCGCTGTCCGCCAAGAGACTGCCAGAGAGAGTGTGGCAGAGAG 297 DB 62 GGGAGCTGGGCAACCTGTCTCTCCAGAGAGCTTCTCAAGAGAGAGTGAACAGAGAGAC 121
OY	298 GACCAGGACCGCTGGAACTGAAATCCCAGACAGAAGAAAGCAGAGATCCCTGGCCTTTC 357 DB 122 CGCCGGGAGCCCCCTGAAGTGAATCCCAAGACAGAGAAAGCAGAGATGTGATCTTTC 181
OY	358 CTGAACCCGACTAGTTGCGGCTCGCAGAAATGCAACCTAAAGGCCGGAACAACAGGGCTGA 417 DB 182 TTGGAACAACACTAGTCCGGCTCGAAGAAATGCTCTTAAGGCCGGAAGGCGGGCTCGC 241
OY	418 AGAGCGATCGCAGCCCATATTATGAAGTTCATCCAGCATGAGCAGAGACGAGCGCAGGCA 477 DB 242 CGAGCTATTGCGAGCCCATATTATGAGTTTCTCTCGGCGAGAGAGATGACACAGCA 301
OY	478 GGTTGGAGACGGGACAGTAGTGGCTGGGGAGAACCGAATCAACAGTCCAGCCCTTCG 537 DB 302 GGTGGTAGAGGGAGCAGTGAAGTGCTGGGGAAGACCAATAATCAACAGCTCCAGCCCTTCG 361
OY	538 CCCTCAACCCGCAATCGGGAGATTATATAGCACCCGGCTGGGCTACTACCTGTAC 597 DB 362 CGCTACGACCGGCACATTGGGGAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTAC 421
OY	598 TGTCAGGTGCATTGTATGATGAGGGGAGAGCTGTCTACCTGAAGCTGGACTTGTGTGAT 657 DB 422 TGCTAGGTGCATTGTATGATGAGGGGAAAGGCTGTCTACCTGAAGCTGGACTTGTGTGAT 481
OY	658 GGTGTGCTGGCCCTGCGCTGCTGGAGGAATTTCTACGACCTACGCGCCAGTTCCCTCGGG 717 DB 482 GGTGTGCTGGCCCTGCGCTGCTGGAGGAATTTCTACGACCTACGCGCCAGCAAGCTCTCTGGG 541
OY	718 CCCGAGTCCGCTCMGCGCAGGTGTGTGGGGGCTGGGCCCTCGGCGCAGAGTCTCTCCG 777 DB 542 CCCGAGTCCGCTTGTGTGCGCAGGTGTGTGGGGTGTGGCCCTCGGCGCAGAGTCTCTCCG 601
OY	778 CGGATCCGACCTCCCTCGGGGCCCATCTCAAGGCTGCCCTTCCTCACTACTTTCGA 837 DB 602 CGGATCCGACCTCCCTCGGGGCTCATCTTAAGGCTGCCCTTCCTCACTACTTTCGA 661
OY	838 CTCCTTCAGGTTCACTGAGGGGCCCTGTCTCCCAACAGTGTCCCAAGGCTGCCGGCTCC 897 DB 662 CTCCTTCAGGTTCACTGAGGGGCCCTGTCTCCCAAGTTCCTTAACCTTCCCTGGCTCC 721
OY	898 -----CTTCGAGAGCTCTGGGCAACCGGCTCCCTTCGCGCCACGCTCAGCGGCTTTC 952 DB 722 AGGAGCATTCACCACTCTCCCTTAAGCCCAACCCCACTCTCCCAACCCCTC-GCTGTCTTC 780
OY	953 GCTTCAGAGCTCCGCTCCCTCTTAAGAGGTGCTGGGCTGTTCACAGTGTTCCTACCC 1012 DB 781 GGTTCAGTCTCTCTCTCC--TCAAAGGAGCGACAGGCTTGTTCACATGTTCATCTCC- 837
OY	1013 ACATAAATPACGTATTTCCCACTCTTAATCTTACAACTCCCAACCGCCACTCTCCACTC 1072 DB 838 -----ACAGAGATPTCTTGTCTTCTTCTTAACATCCCAACCAACAACATCTTCAACCTC 891
OY	1073 ACTAGCTCCCAATCCCTGAGACCTTTTGAAGGCCCAAGTATGTGACACTCCCGCGGGCA 1132 DB 892 ATTAGCTCCCAAGGCCCTAC-----TTATCTCTBACTCTCCCAACCCACT 936
OY	1133 CAGACCCCGAGGCAATTTGTTCACTGTACTGTGTGGGCAAGATGGGTCACAGAGACC 1192

XX OS Homo sapiens.
XX PN W09514772-A1.
XX PD 01-JUN-1995.
XX PF 11-NOV-1994; 94WO-JP01916.
XX PR 12-NOV-1993; 93JP-0355504.
XX PA (MATS/) MATSUBARA K.
XX PA (OKUB/) OKUBO K.
XX PI Matsubara K, Okubo K;
XX DR WPI: 1995-206931/27.
XX XX
XX PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
XX PT for diagnosis of abnormal cell function, by preparing cDNA that
XX PT reflects relative abundance of corresp. mRNA in specific human
XX PT tissues
XX PS Claim 1; Page 1067; 2245pp; Japanese.
XX XX
XX CC A single-stranded DNA (or its complementary strand or the corresp.
XX CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
XX CC given in AAT19001-726837 and which is able to hybridise to part of
XX CC human genomic DNA, cDNA or mRNA is claimed. The GS (gene Signature)
XX CC sequences were obtained from 3'-directed cDNA libraries prepared
XX CC from various human tissues; synthesis of cDNA was initiated from the
XX CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
XX CC untranslated sequence is unique to a particular mRNA species, almost
XX CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
XX CC is constructed so as to reflect accurately the relative abundance of
XX CC different mRNAs in the particular tissue from which it was derived.
XX CC The appearance frequency of a given GS in a cDNA library can be
XX CC determined (esp. using primers and probes derived from the GS
XX CC sequences) as a means of diagnosing abnormal cell function or for
XX CC recognising different cell types.
XX S0 Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 5 other;
Query Match 17.9%; Score 245.4; DB 16; Length 282;
Best Local Similarity 97.3%; Pred. No. 5.3e-43;
Matches 257; Conservative 0; Mismatches 6; Indels 1; Gaps 1;
OY 1111 GATCTGACCTCCCTGGCCACAGACCCCGAGGCGATTGTTCACGTACTCTGTGGG 1170
DB 1 GATCTGACCTCCCTGGCCACAGACCCCGAGGCGATTGTTCACGTACTCTGTGGG 60
OY 1171 CAAGGATGGTCCAGAAACCCCACTTCAGGACCTAAGAGGGCTGGACCTG-GGGGCAG 1229
DB 61 CAAGGATGGTCCAGAAACCCCACTTCAGGACCTAAGAGGGCTGGACCTGGGGAG 120
OY 1230 GAAGCCAAGAGACTGGCTAGGCGCAGAGATTCCCAATGATGAGGGCGAGAAACAGA 1289
DB 121 GAAGCCAAGAGACTGGCTAGGCGCAGAGATTCCCAATGATGAGGGCGAGAAACAGA 180
OY 1290 CAAGCTCTCCCTTGAGAAATTCCTGTGGATTTTAAAAAGATATATATTTATATTA 1349
DB 181 CAAGCTCTCCCTTGAGAAATTCCTGTGGATTTTAAAAAGATATATATTTTNNATTA 240
OY 1350 TTGTGACAAATGTTGATAATGG 1373
DB 241 TTGTGACAAATGTTGNTAATGG 264

RESULT 13
ID ABR29540 standard; cDNA: 195 BP.
XX AC ABR29540;

XX XX
XX DT 23-APR-2002 (first entry)
XX DE Colon adenocarcinoma-specific cDNA #66.
XX KW Human; colon adenocarcinoma; colon cancer; tumour; gene; ss.
XX OS Homo sapiens.
XX PN W0200196389-A2.
XX PD 20-DEC-2001.
XX PF 07-JUN-2001; 2001WO-US18574.
XX PR 09-JUN-2000; 2000US-210667P.
XX PR 22-NOV-2000; 2000US-252614P.
XX PA (CORI-) CORIXA CORP.
XX PI Meagher MJ, King GF, Xu J, Secrett H;
XX DR WPI: 2002-098052/13.
XX XX
XX PT New isolated polynucleotide encoding a polypeptide comprising a portion
XX PT of colon tumour protein, for detection, diagnosis and therapy of human
XX PT colon cancer -
XX PS Claim 1; Page 133; 21pp; English.
XX XX
XX CC The invention relates to an isolated polynucleotide (I) encoding a
XX CC polypeptide (II) comprising a portion of a colon tumour protein. A new
XX CC oligonucleotide (III) that hybridises to (I) is useful for antigen
XX CC determining the presence of a cancer in a patient. (II) or antigen
XX CC presenting cells expressing (I) are useful for stimulating and/or
XX CC expanding T cells specific for a tumour protein, by contacting T cells
XX CC with (I), (II) or antigen-presenting cells that express (I), (II),
XX CC or antigen presenting cells that express (II) are useful for treating
XX CC colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated
XX CC from a patient with (I), (II), or antigen presenting cells that express
XX CC (II), so that T cells proliferate, and administering to the patient an
XX CC effective amount of the proliferated T cells, thus inhibiting the
XX CC development of a cancer in the patient. A new composition is useful for
XX CC stimulating an immune response in a patient. (I) or (II) is useful in
XX CC vaccines and pharmaceutical compositions for prevention and treatment of
XX CC colon cancer and for the diagnosis and monitoring of the cancers. (I),
XX CC (II) or an antibody against (II) is useful for detection, diagnosis and/
XX CC or therapy of human colon cancer. (I) is useful as a probe or primer for
XX CC nucleic acid hybridisation, and in the design and preparation of ribozyme
XX CC molecules for inhibiting expression of (II) in tumour cells. ABR29475-
XX CC ABR29491 represent human colon adenocarcinoma-specific cDNA sequences of
XX CC the invention.
XX S0 Sequence 195 BP; 49 A; 51 C; 58 G; 37 T; 0 other;
Query Match 14.1%; Score 193.4; DB 24; Length 195;
Best Local Similarity 99.5%; Pred. No. 5.7e-32;
Matches 194; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 1130 CCACGAGCCCCCAGGCGATTGTGTTCACGTACTCTGTGGCGAAGATGGTCCAGAGA 1189
DB 1 CCACGAGCCCCCAGGCGATTGTGTTCACGTACTCTGTGGCGAAGATGGTCCAGAGA 60
OY 1190 CCCCACCTTCAGGCACTAAGAGAGGGCTGACCTGGCGGCAAGCAAGAGACTGGGCC 1249
DB 61 CCCCACCTTCAGGCACTAAGAGAGGGCTGACCTGGCGGCAAGCAAGAGACTGGGCC 120
OY 1250 TAGGCGAGAGTTCCCAATGATGAGGGCGAGAAACAGACAAAGCTCCTCCCTTGAGAT 1309
DB 121 TAGGCGAGAGTTCCCAATGATGAGGGCGAGAAACAGACAAAGCTCCTCCCTTGAGAT 180
OY 1310 TCCTGTGGATTTT 1324
XXXXXXXXXXXXXXXXXXXX

Db	181	TCOCGTGATTTT	195
RESULT 14			
AA53491/c			
ID	AA53491	standard: DNA: 114955 BP.	
XX	AA53491:		
XX	05-JUL-1999	(first entry)	
DE	Human adenosine A1 receptor antisense oligonucleotide fragment.		
XX			
KM	Antisense oligonucleotide; multiple target; antisense treatment;		
KM	impaired respiration; inflammation; lung disease;		
KM	pulmonary vasoconstriction; inflammation; allergic rhinitis;		
KM	acute asthma; allergy; asthma; impeded respiration;		
KM	respiratory distress syndrome; pain; cystic fibrosis;		
KM	pulmonary hypertension; pulmonary vasoconstriction; emphysema;		
KM	chronic obstructive pulmonary disease; leukemia; carcinoma;		
KM	colon cancer; breast cancer; lung cancer; pancreatic cancer;		
KM	hepatocellular carcinoma; kidney cancer; melanoma; hepatic metastasis;		
KM	prostate cancer; ss.		
XX			
OS	Synthetic.		
XX			
PN	MO9913868-A1.		
XX			
PD	25-MAR-1999.		
XX			
PF	17-SEP-1998; 98WO-US19419.		
XX			
PR	09-JUN-1998; 98US-0093972.		
PR	17-SEP-1997; 97US-0059160.		
XX			
PA	(UYEC-) UNIV EAST CAROLINA.		
XX			
P1	Nyce JW;		
XX			
DR	WPI: 1999-229400/19.		
XX			
PT	New antisense oligonucleotides used in treatment of, e.g. pulmonary		
PT	vasoconstriction		
XX			
PS	Disclosure: Page 37; 120pp; English.		
XX			
CC	The specification describes antisense oligonucleotides (AA52869-X55271)		
CC	directed against at least 2 mRNAs selected from target genes, coding and		
CC	non-coding regions of RNAs corresponding to target genes, gene		
CC	initiation codons, genomic flanking regions, Intron-exon borders, the		
CC	5'-end, the 3'-end and the junction between coding and non-coding		
CC	regions and all segments of RNAs encoding proteins associated with one		
CC	or more diseases, conditions or mixtures. The antisense oligonucleotides		
CC	may be derived from sequences AA55272-74. These multiple target		
CC	oligonucleotides (specifically AA55180-271) can be used for the		
CC	antisense treatment of diseases and conditions. Typical diseases and		
CC	conditions are those associated with impaired respiration and		
CC	inflammation, including lung diseases, pulmonary vasoconstriction,		
CC	inflammation, allergic rhinitis, acute asthma, allergies, asthma, impeded		
CC	respiration, respiratory distress syndrome, pain, cystic fibrosis,		
CC	pulmonary hypertension, pulmonary vasoconstriction, emphysema, chronic		
CC	obstructive pulmonary disease (COPD), and cancers such as leukemias,		
CC	lymphomas, carcinomas e.g. colon cancer, breast cancer, lung cancer,		
CC	pancreatic cancer, hepatocellular carcinoma, kidney cancer, melanoma,		
CC	hepatic metastases, as well as all types of cancers which may metastasize		
CC	or have metastasized to the lungs, including breast and prostate cancer.		
XX			
SO	Sequence 114955 BP; 6071 A; 29417 C; 36712 G; 21328 T; 21427 other;		
Query Match	5.1%; Score 70.4; DB 20; Length 114955;		
Best Local Similarity	32.7%; Pred. No. 3.6e-05;		
Matches 182; Conservative	58; Mismatches 316; Indels 0; Gaps 0		

Oy	26	CCCCCGGGGCTCCCTCCCGAGATCCCTTCGAGTCCCGGAATGCGGGGCGCGGTGAGGC	85
Db	105272	CCGGCCCGGCGCGCGCANNHHNNNSCGGCCCGGCGCGCGCGCGCCANNHHNNNSCGGCCGC	105213
Oy	86	AAGCACACGCCCCCGCCCCCATGCGCCGCGCTGCGAGACCAGAAGCGAGCGGCGCCGG	145
Db	105212	CGGGGCGCGCCGCGCANNHHNNNSCGGCCGCGCGCGCGCGCGCCGCVNHHNNNSCGGCCGCC	105153
Oy	146	GGAAGCGGGGACCGGCGCTGCTGTCCCGCTCGCGGCTGCGGCGCTGAGGCGCTGAGGACT	205
Db	105132	GCGCGCGCGCGCCGCVNHHNNNSCGGCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG	105093
Oy	206	GCTTCGCGCTCTCTGTCGCGCGCTGTCAATTGGAGAACCGGCGCATGCTGTCCGACAG	265
Db	105092	GCG	105033
Oy	266	AGCCTCGCCGACGAGAGCTGTGTGCGACAGAGAGACAGACGCGCGCTGGAGACTGAATCCC	325
Db	105032	CGGGCG	104973
Oy	326	AGACAGAGAAGACGACGATCTCGCTTTCTGAAACGACTAGTTGCGGCTCGCAGA	385
Db	104972	CCVGNHHNNNSCGGCG	104913
Oy	386	GTGCACCTAAAGCGCGGAAAAACAGCGGCTGCAAGAGCGATGCGACGCCATTATGAATTC	445
Db	104912	CGGCG	104853
Oy	446	ATCCACGACTGGACAGGACGCGACGACGAGCGAGTGTCGACGGGACGACTGATGCTGGG	505
Db	104852	VGGCGCGCGGANNHHNNNSCG	104793
Oy	506	AGGAAGCAGCAATCAACAGCTCCAGCCCTTGCGCTACAACCGCCAGATCGGGGAGTTTA	565
Db	104792	VGGCGCGCGGANNHHNNNSCG	104733
Oy	566	TAGTCACCGCGGCTGG 581'	
Db	104732	NSCGCGCGCGCGCGCG 104717	
<hr/>			
RESULT 15			
AAZ17263			
ID	AAZ17263 standard; cDNA, 1337 BP.		
XX	AAZ17263;		
XX			
DT	12-OCT-1999 (first entry)		
XX			
DE	Human gene expression product cDNA sequence SEQ ID NO:4735.		
XX			
KW	Human; gene; gene expression product; diagnosis; therapy; probe; detection; mapping; tissue typing; profiling; forensic; cancer; genetic analysis; colorectal cancer; breast cancer; lung cancer; ss.		
OS	Homo sapiens.		
XX			
PN	M09938972-A2.		
XX			
PD	05-AUG-1999.		
XX			
PF	28-JAN-1999; 99MO-US01619.		
XX			
PR	03-APR-1998; 98US-0080666. 28-JAN-1998; 98US-0072910. PR 24-FEB-1998; 98US-0075954. PR 31-MAR-1998; 98US-0080114. PR 03-APR-1998; 98US-0080515.		
XX			
PA	(CHIR) CHIRON CORP.		
XX	(HYSE-) HYSED INC.		
IX	Crkvenjakov R, Dickson M, Drmanac R, Drmanac S;		

Search completed: March 31, 2003, 08:17:36
Job time : 437.06 secs

PI Escobedo J, Garcia PD, Garcia V, Giese K, Innis MA;
PI Jones WL, Kassam A, Kennedy GC, Kita D, Labat I;
PI Lanson G, Leshkowitz D, Pot D, Randazzo F, Reinhard C;
PI Stache-Crain B, Sudduth-Klinger J, Williams LT;
DR WPI: 1999-494092/41.

XX Novel human genes and their expression products which are
PT differentially expressed in different cell types
XX

PS Claim 1; Page 2250-2251; 2479pp; English.

XX The present invention describes a library of human polynucleotides
CC comprising the sequences given in AA212532 to AA217779. Also described is
CC a method of detecting differentially expressed genes correlated with the
CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one
CC of the 5248 polynucleotide sequences given in AA212532 to AA217779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensics, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists.
XX

SO Sequence 1337 BP; 42 A; 577 C; 27 G; 22 T; 669 other;

Query Match 4.4%; Score 61; DB 20; Length 1337;
Best Local Similarity 32.1%; Pred. No. 0.0013;

Matches 145; Conservative 0; Mismatches 307; Indels 0; Gaps 0;

OY 690 CTCAGCCACTGGGCGGCTTCCCTGGGGCCCAAGTCCGCTCTGCGAGGTCTGGGCT 749
DB 382 CANNNNCCNN 441
OY 750 GTTGACCTGAGGAGGCTCTCCCTGGGATCCGACCTCCCTGGGCGCATCTCAA 809
DB 442 CCNCCGCCNNCC 501
OY 810 GGGTCCCCCTCTCTCACTTCTGAGCTCTTCCAGGTTCTAGAGGGCCCTGGTCTC 869
DB 502 NCNCCGCCGCCGCCGCCGCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCC 561
OY 870 CCCACAGTGTGCCAGGCTGCCGGCTCCCTGACAGCTCTGTGGGACCCGGTCCCTC 929
DB 562 CCCNCCNNCC 621
OY 930 TGCCCCACCTGAGCGGCTTTGCTCCAGACCTCCCTCCCTAGAGGCTGCTGGG 989
DB 622 CCNCCNNCCGCCGCCGCCGCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNCC 681
OY 990 CCGTTCACGTGTTTTCATCCACATAAATACAGTATTCGACCTTATCTTACAACTC 1049
DB 682 NNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCC 741
OY 1050 CCCCACGGCCACTCTCCACCTAGTCCCAATCCCTGACCCCTTGGAGCCCCCAG 1109
DB 742 CCCCNCNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCC 801
OY 1110 TGATCTGACTGCCCTGGCCAGACGCC 1141
DB 802 NNNNNNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCCNNCC 833

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2003, 07:58:15 ; Search time 4329.73 Seconds

(without alignments)
9228.782 Million cell updates/sec

Title: US-09-245-198A-3

Perfect score: 1373
Sequence: 1 atgtcatgtttagacttga.....gacaaatgttgataaatg 1373

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

```
Database :
1: GenEmbl:*
2: gb_ba:*
3: gb_hcg:*
4: gb_in:*
5: gb_om:*
6: gb_ov:*
7: gb_pat:*
8: gb_ph:*
9: gb_pl:*
10: gb_pr:*
11: gb_ro:*
12: gb_sts:*
13: gb_sy:*
14: gb_un:*
15: gb_vl:*
16: em_ba:*
17: em_fun:*
18: em_hum:*
19: em_in:*
20: em_mu:*
21: em_om:*
22: em_or:*
23: em_ov:*
24: em_pat:*
25: em_ph:*
26: em_pl:*
27: em_ro:*
28: em_sts:*
29: em_un:*
30: em_vl:*
31: em_hcg_hum:*
32: em_hcg_inv:*
33: em_hcg_other:*
34: em_hcg_mus:*
35: em_hcg_pln:*
36: em_hcg_rtd:*
37: em_hcg_mam:*
38: em_hcg_vit:*
39: em_sy:*
40: em_hgtgo_hum:*
41: em_hgtgo_mus:*
42: em_hgtgo_other:*
```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1320.2	96.2	1353	6	AX201324 Sequence
2	1320.2	96.2	1368	9	AF055872 Homo sapi
3	1285	93.6	1306	9	AF030099 Homo sapi
4	1226.4	89.3	1236	6	AR140407 Sequence
5	1096.8	79.9	1651	9	BC019047 Homo sapi
6	768.2	56.0	177703	2	AC016876 Homo sapi
7	754	54.9	215795	2	AC127470 Pan trogl
8	629.2	45.8	898	6	AX180714 Sequence
9	614.6	44.8	1168	10	AF030100 Mus muscu
10	364.2	26.5	153553	2	AC126921 Bos tauru
11	309.6	22.5	203083	2	AC069459 Mus muscu
12	309.6	22.5	234182	10	AL603707 Mouse DNA
13	254.2	18.5	138792	2	AC119115 Rattus no
14	246.4	17.9	177555	2	AC130192 Sus scrofa
15	242.4	17.7	161428	2	AC126825 Canis fam
16	193.4	14.1	195	6	AX379024 Sequence
17	158.6	11.6	184026	2	AC098923 Rattus no
18	111	8.1	203281	2	AC126237 Canis fam
19	100.6	7.3	139405	2	AC126239 Felis cat
20	87.2	6.4	7218	6	I66494 Sequence 14
21	77.6	5.7	125020	9	AF429315 Homo sapi
22	74.8	5.4	303091	2	AC084799 Mus muscu
23	71	5.2	121451	2	AC111931 Rattus no
24	70.8	5.2	73948	2	AC022556 Homo sapi
25	70.6	5.1	144979	2	AC016280 Homo sapi
26	70.4	5.1	220469	2	AC074307 Mus muscu
27	70.2	5.1	298166	2	AC087563 Homo sapi
28	69.2	5.0	936	8	CNS01AYK Botrytis
29	69.2	5.0	82897	2	AC022595 Homo sapi
30	68.2	5.0	110737	2	AC011105 Homo sapi
31	68.2	5.0	265537	2	AC087228 Mus muscu
32	68	5.0	181344	2	AC125880 Rattus no
33	68	5.0	300695	2	AC079431 Mus muscu
34	67.6	4.9	626	8	CNS01BRP AL110716 Botrytis
35	67.2	4.9	66624	2	AC101520 Mus muscu
36	67.2	4.9	147124	2	AC022768 Homo sapi
37	67.2	4.9	177883	2	AC046159 Homo sapi
38	67.2	4.9	181988	2	AC090552 Homo sapi
39	67	4.9	74138	2	AC021272 Homo sapi
40	66.8	4.9	62649	2	AC022552 Homo sapi
41	66.8	4.9	73628	2	AC021793 Homo sapi
42	66.8	4.9	100511	2	AC010774 Homo sapi
43	66.6	4.9	167830	2	AC103026 Rattus no
44	66.4	4.8	840	8	CNS01BNS AL114464 Botrytis
45	66.2	4.8	61834	2	AC068105 Homo sapi

ALIGNMENTS

```
RESULT 1
LOCUS AX201324 1353 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 3 from Patent WO0153486.
ACCESSION AX201324
VERSION AX201324.1 GI:15391154
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 1353)
Ashkenazi,A.J., Goddard,A., Godowski,P.J., Gurey,A.L.,
Hillan,K.J., Masters,S.A., Pan,J., Pitti,R.M., Roy,M.A., Smith,V.,
Stone,D.M., Watanabe,C.K. and Wood,W.I.
```



```
/product="Ap03/DR3_ligand"
/protein_id="AAC39724.1"
/db_xref="GI:3108231"
/translation="MAARRSORRRGRGEPETALLVPLALGLIALACLILLAVSL
GSRASLSAEPDELVAEDDPSELPDPEESQDDPPELNLRIYPRSRAPKRTK
ARRIAHAYEHPRPQDQAGVGVNVSVEEARINSSPLNRQIGEFIVRAGL
YYLCOVHFDEGKAVYLLKDLIVDGLALRGLIEFEFSAATASISLPOLRLCOVSLAL
RPSLSLRRLPMHLKMAFLPYEGLFOVH"
BASE COUNT      272 a      443 c      389 g      264 t
ORIGIN
Query Match      96.2%; Score 1320.2; DB 9; Length 1368;
Best Local Similarity 99.8%; Pred. No. 7.9e-231;
Matches 1322; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```

Db      841 GCGGCTCCCTCGACAGCTCTGTGGGACACCGGTCCTCCCTGCCCAACCTCAGCCGCT 900
Qy      949 CTTTCTCTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCGGGCTTTACAGTGTTCGA 1008
Db      901 CTTTCTCTCCAGACCTGCCCCCTCCCTCTAGAGGCTGCGGGCTTTACAGTGTTCGA 960
Qy      1009 TCCCAATTAATATACAGTATTCACACTCTATCTACACTCCCGCCAGCCACTCTCCA 1068
Db      961 TCCCAATTAATATACAGTATTCACACTCTATCTACACTCCCGCCAGCCACTCTCCA 1020
Qy      1069 CCTCACTAGCTCCCAATCCTGACCTCTTGAAGGCCCCCAAGTATCTGACTCCCTCTG 1128
Db      1021 CCTCACTAGCTCCCAATCCTGACCTCTTGAAGGCCCCCAAGTATCTGACTCCCTCTG 1080
Qy      1129 GCCACAGACCCCGAGGATGTGTCACTGTACTGTGGGCAAGAGTGGGTCCAGAG 1188
Db      1081 GCCACAGACCCCGAGGATGTGTCACTGTACTGTGGGCAAGAGTGGGTCCAGAG 1140
Qy      1189 ACCCACTTTCAGAGCACTAAGAGGGGCTGACCTGCGCGAGAGCAACCAAGACTGGGC 1248
Db      1141 ACCCACTTTCAGAGCACTAAGAGGGGCTGACCTGCGCGAGAGCAACCAAGACTGGGC 1200
Qy      1249 CTAGGCCAGAGGTTCCCAATGTAGAGGGCGAGAAACAAGACACTCTCTCTTGAGAA 1308
Db      1201 CTAGGCCAGAGGTTCCCAATGTAGAGGGCGAGAAACAAGACACTCTCTCTTGAGAA 1260
Qy      1309 TTCCCTGTGATTTTAAACGATATATTTTATATATGACAAATGTGTATA 1368
Db      1261 TTCCCTGTGATTTTAAACGATATATTTTATATATGACAAATGTGTATA 1320
Qy      1369 AATGG 1373
Db      1321 AATGG 1325

RESULT 3
AF030099      1306 bp      mRNA      linear      PRI 20-DEC-1997
LOCUS      AF030099
DEFINITION      Homo sapiens TWEAK mRNA, complete cds.
ACCESSION      AF030099
VERSION      AF030099.1 GI:2707218
KEYWORDS
SOURCE      Homo sapiens.
ORGANISM      Homo sapiens.
REFERENCE      1 (bases 1 to 1306)
AUTHORS      Chicheportliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H.,
Hession,C., Garcia,I. and Browning,J.L.
TWEAK, a new secreted ligand in the tumor necrosis factor family
that weakly induces apoptosis
J Biol. Chem. 272 (51), 32401-32410 (1997)
JOURNAL      JOURNAL
MEDLINE      98070415
PUBMED      9405449
REFERENCE      2 (bases 1 to 1306)
AUTHORS      Bourdon,P., Hession,C., Tizard,R. and Browning,J.
TITLE      Direct Submission
JOURNAL      Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center,
Cambridge, MA 02142, USA
FEATURES
source
1..1306
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
/tissue-type="tonsil"
18..767
/note="Ligand in the TNF family; secreted protein; start
codon not verified experimentally"
/product="TWEAK"
/protein_id="AAC39724.1"
/db_xref="GI:2707219"
```

```

/translation="MAARRSRRRGRREPGTALLVPLALGLALACTLLAAVSL
GSRASLSADPEAQEELVAEEDDPSELNQTEESODPAFLRLRPRRSARFKRTR
ARRAIAAHYEVRPRGDDAAGVDTGATVPRNNSPFLRYNNOIEFLVTRAGL
XYLCOVHDEBKAKYIKIKLIDLVGIALRCHLEERSATMASSLGQRLRCQVSGLLAL
RGSRLRITTEMAHKAPFLITTEFLPQVH"
BASE COUNT      247 a      434 c      368 g      257 t
ORIGIN
Query Match      93.6%; Score 1285; DB 9; Length 1306;
Best Local Similarity 100.0%; Pred. No. 2,1e-224;
Matches 1285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 89 CACAGCCCCCCCCCATGCGCCCGCTGAGAGCCAGAGCGGAGGGGCGCGGGGG 148
   |||||
Db 1 CACAGCGCCCCCGCCCATGCGCCCGCTGAGAGCCAGAGCGGAGGGGCGCGGGGG 60
QY 149 AGCGGGGACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 208
   |||||
Db 61 AGCGGGGACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
QY 209 TCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 268
   |||||
Db 121 TCGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180
QY 269 CTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 328
   |||||
Db 181 CTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 329 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
   |||||
Db 241 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 389 CACCTAAAGGCGGAAACACGCGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 448
   |||||
Db 301 CACCTAAAGGCGGAAACACGCGGCTCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
QY 449 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 508
   |||||
Db 361 CAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
QY 509 AAGCCAGAAATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 568
   |||||
Db 421 AAGCCAGAAATCAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 569 TCAACCGGGGCTGAGCTCTACTACTACTACTACTACTACTACTACTACTACTACTACT 628
   |||||
Db 481 TCAACCGGGGCTGAGCTCTACTACTACTACTACTACTACTACTACTACTACTACTACT 540
QY 629 TCTACTGGAAGCTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 688
   |||||
Db 541 TCTACTGGAAGCTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 689 TCTCAGCAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 748
   |||||
Db 601 TCTCAGCAGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY 749 TGTGTGGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 808
   |||||
Db 661 TGTGTGGGCTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 720
QY 809 AGGCTGCGCCCTTCTCCTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 868
   |||||
Db 721 AGGCTGCGCCCTTCTCCTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 780
QY 869 CCCCAGAGTGTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 928
   |||||
Db 781 CCCCAGAGTGTGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 840
QY 929 CTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 988
   |||||
Db 841 CTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
QY 989 GCCTGTTCACGCTTTTCCATCCCATTAATACAGTATTCCACACTTATTCTTACACT 1048

```

```

|||||
Db 901 GCCTGTTCACGCTTTTCCATCCCATTAATACAGTATTCCACACTTATTCTTACACT 960
QY 1049 CCCCCACCGCCCACTCCACACTCAGTACTGCTCCCAATCCGAGCCTTGGAGGCCCA 1108
   |||||
Db 961 CCCCCACCGCCCACTCCACACTCAGTACTGCTCCCAATCCGAGCCTTGGAGGCCCA 1020
QY 1109 GTGATCTGACATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1168
   |||||
Db 1021 GTGATCTGACATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
QY 1169 GGCAGAGATGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1228
   |||||
Db 1081 GGCAGAGATGCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1140
QY 1229 GGAAGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1288
   |||||
Db 1141 GGAAGCCAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
QY 1289 ACAAGCTCTCTCTCTGAGAGATTCCTGCTGATTTTAAACAGATTTATTTATTTATTT 1348
   |||||
Db 1201 ACAAGCTCTCTCTCTGAGAGATTCCTGCTGATTTTAAACAGATTTATTTATTTATTT 1260
QY 1349 ATTGTGACAAATGTTGATTAATGG 1373
   |||||
Db 1261 ATTGTGACAAATGTTGATTAATGG 1285
RESULT 4
LOCUS      AR140407      1236 bp      DNA      linear      PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6207642.
ACCESSION  AR140407
VERSION     AR140407.1  GI:14482903
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
REFERENCE   1 (bases 1 to 1236)
AUTHORS     Wiley,S.R.
TITLE       Member of the TNF family useful for treatment and diagnosis of
            disease
            Patent: US 6207642-A 1 27-MAR-2001;
FEATURES
            source      Location/Qualifiers
            source      1..1236
                        /organism="unknown"
BASE COUNT      225 a      416 c      358 g      237 t
ORIGIN
Query Match      89.3%; Score 1226.4; DB 6; Length 1236;
Best Local Similarity 99.5%; Pred. No. 1e-213;
Matches 1230; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
QY 106 ATGGCGGCGGCTGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 165
   |||||
Db 1 ATGGCGGCGGCTGAGAGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 166 CTGCTCCGCTGCGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 225
   |||||
Db 61 CTGCTCCGCTGCGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGGCTGAG 120
QY 226 GTGCTCAGTTTGGGAGAGCGGCGATGCTGCTGCGCCAGAGAGCTTCCAGAGAGAGCTG 285
   |||||
Db 121 GTGCTCAGTTTGGGAGAGCGGCGATGCTGCTGCGCCAGAGAGCTTCCAGAGAGAGCTG 180
QY 286 GTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 345
   |||||
Db 181 GTGGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 346 CTGCGGCTTTCTGAAACGAGTGTGCGCTCGGAGAGAGTGCACCTAAAGGCGGAGAA 405
   |||||
Db 241 CTGCGGCTTTCTGAAACGAGTGTGCGCTCGGAGAGAGTGCACCTAAAGGCGGAGAA 300

```

```

Oy 406 ACACGGGCTCGAAGAGCGATCGACGCCCATATGAAATTTCATCCAGACCTGAGACAG 465
    |||||||
Db 301 ACAGGGGCTCGAAGAGCGATCGACGCCCATATGAAATTTCATCCAGACCTGAGACAG 360
Oy 466 GAGAGCGCAGGAGGCTGTGAGCGGACAGTGAAGTGGCTGGGAGGAAGCCAGAAATCAACAGC 525
    |||||||
Db 361 GAGAGCGCAGGAGGCTGTGAGCGGACAGTGAAGTGGCTGGGAGGAAGCCAGAAATCAACAGC 420
Oy 526 TCCAGACCCCTGCGGTACACACCGGCAGATCGGGGAGTTATATGTACACCCGGGCTGGGCTC 585
    |||||||
Db 421 TCCAGACCCCTGCGGTACACACCGGCAGATCGGGGAGTTATATGTACACCCGGGCTGGGCTC 480
Oy 586 TACTACTGTACTGTACAGTGCATTTGATGAGGGGAAAGGCTGTACTTACCTGAAGCTGAC 645
    |||||||
Db 481 TACTACTGTACTGTACAGTGCATTTGATGAGGGGAAAGGCTGTACTTACCTGAAGCTGAC 540
Oy 646 TTGCTGTGTGATGTGTGTGCTGGCCCTGGGCTGCTGGAGGAATTTCTACACCTGGGCGC 705
    |||||||
Db 541 TTGCTGTGTGATGTGTGTGCTGGCCCTGGGCTGCTGGAGGAATTTCTACACCTGGGCGC 600
Oy 706 ACTTCCCTCGGGGCGCCAGTCCGGCTGCGGAGGTGTGGGGCTGTGGCCCTGGGCGCA 765
    |||||||
Db 601 ACTTCCCTCGGGGCGCCAGTCCGGCTGCGGAGGTGTGGGGCTGTGGCCCTGGGCGCA 660
Oy 766 GGGTCCCTCGGAGTCCGACACCTCCCTGCGGCGCATCTCAAGGCTGCCCTTCTCTC 825
    |||||||
Db 661 GGGTCCCTCGGAGTCCGACACCTCCCTGCGGCGCATCTCAAGGCTGCCCTTCTCTC 720
Oy 826 ACCTACTTGGACTTTTCCAGGTTACTGAGGGGCGCTGGTCTCCACACAGTGTCCAG 885
    |||||||
Db 721 ACCTACTTGGACTTTTCCAGGTTACTGAGGGGCGCTGGTCTCCACACAGTGTCCAG 780
Oy 886 GGTGCGGGTCCCGTCGACGCTCTGGGCGACCGGTCGCCCTGTCGCCACCTCAGCC 945
    |||||||
Db 781 GGTGCGGGTCCCGTCGACGCTCTGGGCGACCGGTCGCCCTGTCGCCACCTCAGCC 840
Oy 946 GCTCTTGTGCTCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTGTTCAAGTGTCTT 1005
    |||||||
Db 841 GCTCTTGTGCTCAGACCTGCCCCCTCTCTAGAGGCTGCTGGGCTGTTCAAGTGTCTT 900
Oy 1006 CCATCCCAATTAATACAGTATTCCTTATTTTACACTCCGCCACCGCCCACTCT 1065
    |||||||
Db 901 CCATCCCAATTAATACAGTATTCCTTATTTTACACTCCGCCACCGCCCACTCT 960
Oy 1066 CCACCTACTAGCTCCCAATCCCGACCTTGAAGGCGCCGAGTGTGACTGCCCTCC 1125
    |||||||
Db 961 CCACCTACTAGCTCCCAATCCCGACCTTGAAGGCGCCGAGTGTGACTGCCCTCC 1020
Oy 1126 CTGGCCAGACAGACCCCGAGGCAATGTGTCTACTGTACTTGTGGCAAGATGGGTCCAG 1185
    |||||||
Db 1021 CTGGCCAGACAGACCCCGAGGCAATGTGTCTACTGTACTTGTGGCAAGATGGGTCCAG 1080
Oy 1186 AAGACCCCACTTCAAGCACTAAAGAGGGCTGAGCTGGCGGAGAGCAAGCAAGACTG 1245
    |||||||
Db 1081 AAGACCCCACTTCAAGCACTAAAGAGGGCTGAGCTGGCGGAGAGCAAGCAAGACTG 1140
Oy 1246 GGGCCAGAGCCAGAGATTTCCCAATGTGAGGGCGAGAAAAGACAGAGTCTCTCCCTGA 1305
    |||||||
Db 1141 GGGCCAGAGCCAGAGATTTCCCAATGTGAGGGCGAGAAAAGACAGAGTCTCTCCCTGA 1200
Oy 1306 GAATTCCTGTGATTTTAAACAGATATTTATTT 1341
    |||||||
Db 1201 GAATTCCTGTGATTTTAAACAGATATTTATTT 1236

```

```

RESULT 5
BC019047      1651 bp      mRNA      linear      PRI 11-DEC-2001
LOCUS        Homo sapiens, similar to tumor necrosis factor (ligand)
DEFINITION   superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
ACCESSION    BC019047
VERSION      BC019047.1 GI:17512138

```

KEYWORDS MGC.

SOURCE Homo sapiens.

ORGANISM Homo sapiens.

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 1651)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-7590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk
Email: cgapsb-remail.nih.gov
Tissue Procurement: Louis Staudt
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc-mgc@nhgri.nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stancirpop, S., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L., H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: ILNL Plate: 30 Row: P Column: 5
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507596.

FEATURES

source

1..1651
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="MGC:20669 IMAGE:4766071"
/tissue_type="Primary B-cells from Tonsils"
/clone_lib="NIH-MGC-48"
/lab_host="DH10B-R"
/note="vector: pOTB7"
106..510
/codon_start=1
/product="Similar to tumor necrosis factor (ligand) superfamily, member 12"
/protein_id="AAH19047.1"
/db_xref="gi:17512139"
/translation="MAARRSQRRRGRGEPSTALLVPLALGLIALCLLAVLSLSRAISAOEPADIELVAEDPDSELPDQEESEODPAFLNLRVRRSPKGRKTRARRAIAHYEHVPRPGDGAQADSDGTYTCLRP"

CDS

BASE COUNT 344 a 517 c 481 g 309 t
ORIGIN

Query Match 79.9%; Score 1096.8; DB 9; Length 1651;
Best Local Similarity 84.8%; Pred. No. 4.1e-190;
Matches 1347; Conservative 0; Mismatches 2; Indels 239; Gaps 1;

```

Oy 25 TCCGCGCGCGGCTCGCCCTCCCGGATCCCTCGGGGTCGCGGAGTGGGGGCGGTGAGG 84
    |||||||
Db 25 TCCGCGCGCGGCTCGCCCTCCCGGATCCCTCGGGGTCGCGGAGTGGGGGCGGTGAGG 84
Oy 85 CAGGCACAGCCCCCGCCCGCATGCGCGCCGTCGAGACAGAGCGAGGCGGCGCCG 144
    |||||||
Db 85 CAGGCACAGCCCCCGCCCGCATGCGCGCCGTCGAGACAGAGCGAGGCGGCGCCG 144
Oy 145 GGGAGCGCGGACCGCCCTGCTGCTCCGCTCGCGCTGCGGCTTGGCCCTGGCGCTGGCC 204
    |||||||
Db 145 GGGAGCGCGGACCGCCCTGCTGCTCCGCTCGCGCTGCGGCTTGGCCCTGGCGCTGGCC 204

```

OY	205	TGCTCGGGCCCTCCGTCGTCGGCCGTCGTCAGTTTGGGAGCGGGGCACTGCTGTCGCCACG	264
Db	205	TGCTCGGGCCCTCCGTCGTCGGCCGTCGTCAGTTTGGGAGCGGGGCACTGCTGTCGCCACG	264
OY	265	GAGCCTGCGCAGGAGGAGCTGTGTGGCAGAGAGAGACACAGAGACCCGTGCGAATCGATCC	324
Db	265	GAGCCTGCGCAGGAGGAGCTGTGTGGCAGAGAGAGACACAGAGACCCGTGCGAATCGATCC	324
OY	325	CAGACAGAGAAACCCAGAGATCCTGCGCCCTTCTCTAACCAGCTAGTTCGGCCTGCAGAG	384
Db	325	CAGACAGAGAAACCCAGAGATCCTGCGCCCTTCTCTAACCAGCTAGTTCGGCCTGCAGAG	384
OY	385	AGTGCACCTTAAGGCGGAAAAACACGGGGCTGSAAGAGCGATGTGCAACCCCATATGAACTT	444
Db	385	AGTGCACCTTAAGGCGGAAAAACACGGGGCTGSAAGAGCGATGTGCAACCCCATATGAACTT	444
OY	445	CATCCACGACCTTGACAGACGAGCGAGCGCAGG	475
Db	445	CATCCACGACCTTGACAGAGGACGAGCGGACGAGATGSAAGTTTACCAACTTGTCTGAGG	504
OY	476	-----	475
Db	505	CCATGAGATCTAAGTGTGGAGCCAGATTTGAACCAGCTAGATGTGCTGTACT	564
OY	476	-----	475
Db	555	GGACATGCTTCGATGAAAGGCCAGGTGCTGCTGCAGGGGTGAGGGGTCCATTCAGAGGG	624
OY	476	-----	475
Db	625	CCACATCCAAAAAGGGGAGAGGGAGTTTCCAGAAAGAAAGAGAGACATCTCCACACA	684
OY	476	-----CAGCTGTGACGGGAGACGTGAGTGCCTGGG	505
Db	685	TTACAGAGGGTCAAAGGAGAGAACAGAGACCAAGTGTGGACGGGAGACGTGAGTGCCTGGG	744
OY	506	AGSAAAGCCAGAAATCAACAGGCTCCAGCCCTGTGGGTCTACAACCGCCGAGTCGGGGAGTTA	565
Db	745	AGGAAAGCCAGAAATCAACAGGCTCCAGCCCTGTGGGTCTACAACCGCCGAGTCGGGGAGTTA	804
OY	566	TAGTCACCCGGGCTGGGGCTGTACTACTCTGACTGTCTCAGGTGCACTTTGATGAGGGGAAAG	625
Db	805	TAGTCACCCGGGCTGGGGCTGTACTACTGTGATGTGTGCTGGGCTCGCTGCTGAGAG	864
OY	626	CTGTCTACCTTAAGCTGAGCTTGTGTGTGATGTGTGCTGCTGCGCTGCTGAGAG	685
Db	865	CTGTCTACCTTAAGCTGAGCTTGTGTGTGATGTGTGCTGCTGCGCTGCTGAGAG	924
OY	686	AATTCCTCAGGCATCGGCGCCAGTTCCCTGGGGCCCGAGCTCCGCTCCGCGAGGTGTG	745
Db	925	AATTCCTCAGGCATCGGCGCGAGTTCCCTCGGGCCCGAGCTCCGCTCTCTCCAGGTGTG	984
OY	746	GCGTGTGTGGCCCTCGGCGCAGGGGTCTCCCTGCGGATCGCAGACCCCTCGGCGCCATC	805
Db	985	GCGTGTGTGGCCCTCGGCGCGAGGGGTCTCCCTGGGATTCGCAACCTCCCTCGGCGCATC	1044
OY	806	TCAAGGCTGCCCCCTTCTCACTACTTGGGACTTTCAGAGTTCACTGAGGGGCCCTG	865
Db	1045	TCAAGGCTGCCCCCTTCTCACTACTTGGGACTTTCAGAGTTCACTGAGGGGCCCTG	1104
OY	866	TCTCCCCCACTGCTGCCAGAGCTGCCGGCTCCCTCGACAGACTCTGTGGGACCCGGTCC	925
Db	1105	TCTCCCCCACTGCTGCCAGAGCTGCCGGCTCCCTCGACAGACTCTGTGGGACCCGGTCC	1164
OY	926	CCCTGCGCCCACTTCAGCGCGCTTTTGGTCAGAGACTGCGCCCTCCCTCTAAGAGCTGC	985
Db	1165	CCCTGCGCCCACTTCAGCGCGCTTTTGGTCAGAGACTGCGCCCTCCCTCTAAGAGCTGC	1224
OY	986	TGGGCTGTTTACAGTGTTTTTCATCCACATTAATACATATTCCACACTTATCTTACA	1045
Db	1225	TGGGCTGTTTACAGTGTTTTTCATCCACATTAATACATATTCCACACTTATCTTACA	1288
OY	1046	ACTCCCCACCGCCACTCTCCACCTTCACTAGCTCCCACTCCCATGCCCTTTTGGAGGCC	1105

Db	1285	ACTCCCCACCGCCACACTCTCCACCTCACTAGCTTCCCAATCCCTGACCCCTTTGAGGCC	1344
Qy	1106	CCAGTGATCTGAGACTCCCCCCTGGCCACAGACCCCCAGAGGCAATTGTTCCTACTGTACTCT	1165
Db	1345	CCAGTGATCTGAGACTCCCCCCTGGCCACAGACCCCCAGAGGCAATTGTTCCTACTGTACTCT	1404
Qy	1166	GTCGGCCAAGGATGGGTCCAGAAAGACCCCACTTCAGGCACTAAAGAGGGCTGGACCTGGCG	1225
Db	1405	GTCGGCCAAGGATGGGTCCAGAAAGACCCCACTTCAGGCACTAAAGAGGGCTGGACCTGGCG	1464
Qy	1226	GCAGGAAGCCAAAGAGACTGGGGCTAGGCCACAGAGATTTCCCAATATGTGAGGGCGGAGAAC	1285
Db	1465	GCAGGAAGCCAAAGAGACTGGGGCTAGGCCACAGAGATTTCCCAATATGTGAGGGCGGAGAAC	1524
Qy	1286	AAGACAAGCTCCTCCCTTGGAGAAATTCCTGTGGATTTTAAACAGATATATATTTTATTTATTT	1345
Db	1525	AAGACAAGCTCCTCCCTTGGAGAAATTCCTGTGGATTTTAAACAGATATATATTTTATTTATTT	1584
Qy	1346	ATTATTTGACAAATGTTGATTAATGG	1373
Db	1585	ATTATTTGACAAATGTTGATTAATGG	1612

LOCUS	AC016876	LOCUS	AC016876	177703	bp	DNA	linear	HTG	06-AUG-2002	
DEFINITION	Homo sapiens clone RP11-168B7, *** SEQUENCING IN PROGRESS ***	DEFINITION	Homo sapiens clone RP11-168B7, *** SEQUENCING IN PROGRESS ***	10						
ACCESSION	AC016876	ACCESSION	AC016876							
VERSION	AC016876.5	VERSION	AC016876.5	GI:21313830						
KEYWORDS	HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.	KEYWORDS	HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN.							
SOURCE	human.	SOURCE	human.							
ORGANISM	Homo sapiens	ORGANISM	Homo sapiens							
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.	REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homindae; Homo.							
AUTHORS	1 (bases 1 to 177703)	AUTHORS	1 (bases 1 to 177703)							
TITLE	Unpublished	TITLE	Unpublished							
JOURNAL	2 (bases 1 to 177703)	JOURNAL	2 (bases 1 to 177703)							
REFERENCE	Barren,B., Nussbaum,C. and Lander,E.	REFERENCE	Barren,B., Nussbaum,C. and Lander,E.							
AUTHORS	Barren,B., Nussbaum,C. and Lander,E.	AUTHORS	Barren,B., Nussbaum,C. and Lander,E.							
TITLE	Homo sapiens chromosome, clone RP11-168B7	TITLE	Homo sapiens chromosome, clone RP11-168B7							
JOURNAL	Unpublished	JOURNAL	Unpublished							
REFERENCE	2 (bases 1 to 177703)	REFERENCE	2 (bases 1 to 177703)							
AUTHORS	Barren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckery,R., Boguslavsky,L., Boukhalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lebecky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., Mcwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,C.N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.	AUTHORS	Barren,B., Linton,L., Nussbaum,C., Lander,E., Allen,N., Anderson,M., Baldwin,J., Barna,N., Beckery,R., Boguslavsky,L., Boukhalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Domino,M., Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J., Lebecky,J., Lieu,C., Locke,K., Macdonald,P., Marquis,N., Mcwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrim,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,C.N., Subramanian,A., Talamas,J., Tesfaye,S., Tirrell,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.							
TITLE	Direct Submission	TITLE	Direct Submission							
JOURNAL	Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA	JOURNAL	Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA							
REFERENCE	3 (bases 1 to 177703)	REFERENCE	3 (bases 1 to 177703)							
AUTHORS	Barren,B., Nussbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Chazaro,B., Choepely,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Linblad-Touh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,K., McCarthy,M., Meldrim,J., Menus,L., Mihovata,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoll,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,R., Schuppack,R., Seaman,S., Severy,P.,	AUTHORS	Barren,B., Nussbaum,C., Lander,E., All,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Chazaro,B., Choepely,Y., Collymore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,M., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Linblad-Touh,K., Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,K., McCarthy,M., Meldrim,J., Menus,L., Mihovata,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicoll,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neil,D., Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,R., Schuppack,R., Seaman,S., Severy,P.,							


```

Db      182 TTGGAACAACCTAGTCGGCTCCGCAAGAGGCTTAAAGGCGGAGGCGGCGCTCCG 241
Oy      418 AGAGCATGCGAGCCCATTTATGAAATTTCACGACCTGGACAGGACGAGCGGCAAGCA 477
Db      242 CGAGCATTTGCGAGCCCATTTATGAGTTTCATCCCTGGCCAGAGAGATGAGACACAAGCA 301
Oy      478 GGTGTGAGCGGAGCACTGAGTGGCTGGAGAGAACCCAGAAATCAAGCTCCAGCCCTCTG 537
Db      302 GGTGTGAGTGGAGCACTGAGTGGCTGGAGAGAACCCAGAAATCAAGCTCCAGCCCTCTG 361
Oy      538 CGGTAAACCCGCGAGATTCGGGGAGTTTATAGTACCCGGGCTGGGCTTACTACTACTGAC 597
Db      362 CGGTAGAGCCGCGAGATTCGGGGAGTTTATAGTACCCGGGCTGGGCTTACTACTACTGAC 421
Oy      598 TGTGAGGTGCACCTTTGATGAGGGAGGAGGCTGTACTGAACTGAGCTTGGCTGGTGGAT 657
Db      422 TGTGAGGTGCACCTTTGATGAGGGAGGAGGCTGTACTGAACTGAGCTTGGCTGGTGGAT 481
Oy      658 GGTGTGCTGGCCCTGGCGCTGCGTGGAGAAATTCACGCCACTGCGGCGCAGTTCCCTCGG 717
Db      482 GGTGTGCTGGCCCTGGCGCTGCGTGGAGAAATTCACGCCACTGCGGCGCAGTTCCCTCGG 541
Oy      718 CCCGACCTCGGCTCTGCCAGGTGTGTGGGCTGTGGCCCTGGCGGCGGCTCTCCCTG 777
Db      542 CCCGACCTCGGCTGTGGCGAGGTGTGTGGGCTGTGGCCCTGGCGGCGGCTCTCCCTG 601
Oy      778 CGGATCCGAGCCCTGCCCTGGGCGCATTCGAAAGGCTGCCGCCCTCCGCACTTCCGGA 837
Db      602 CGGATCCGAGCCCTGCCCTGGGCGCATTCGAAAGGCTGCCGCCCTCCGCACTTCCGGA 661
Oy      838 CTCCTTCAGAGTTCACAGAGGGGCGCTGCTGCCACAGTGTCCAGGCTGCCGCGCTCC 897
Db      662 CTCCTTCAGAGTTCACAGAGGGGCGCTGCTGCCACAGTGTCCAGTTCCTTCCCTGGCTCC 721
Oy      898 -----CCTGACAGCTCTGTGGGCAACCCGCTGCCCTCTGCCACCCCTCAGCCGCTCTT 952
Db      722 AGGAGATCAGCAGCAGCTCTCCCTAGAGGCGCTGGGCGCTGTGACGTTGTTTCCATGCC 1012
Oy      953 GCTTCAGAGCTGCGCCCTCTCCCTAGAGGCGCTGGGCGCTGTGACGTTGTTTCCATGCC 1012
Db      781 GGTCCAGCTCTGTCTCTCC--TCAAAGGCAAGCAGGAGCTTTGACATGTTTCCATTC-- 837
Oy      1013 ACATAAATACAGTATTCCTCAGCTCTTATCTTACAACTGCCCGCAGCTCTCCAGCTTC 1072
Db      838 -----ACAGAGCTATCTTGTGCTCTTCTTAAACATCCATCCACACAATATCCACCTC 891
Oy      1073 ACTAGTCCCGCAATCCCTGACCTTTGAGAGGCCCGCCAGTATCTGACCTCCCTCCGCGCA 1132
Db      892 ACTAGTCCCGCAATCCCTGACCTTTGAGAGGCCCGCCAGTATCTGACCTCCCTCCGCGCA 936
Oy      1133 CAGACCCCGCAGGAGCTGTCTTCACTGTACTCTGTGGGCAAGGATGGGTCCAGAAAGCC 1192
Db      937 CAGCCGACCAAGTATTGACTTTGTGTGAC----- 968
Oy      1193 CACTTCAGGACTAAGAGGGGCTGGACCTGGCGGCGGCAAGAACCAAGAGACTGGGCTAG 1252
Db      969 -----CAGGCACTGAGATGGGCTGGAGCTGTGGGCAAGGACCAAGCTTGGGACTAG 1023
Oy      1253 GCCAGAGTTCCTCAATGTGAGGGGCGAGAG--AACAGACAGAGCTCCCTTGAAGATTTC 1311
Db      1024 GCCAGAGTTCCTCAATGTGAGGGGCGAGAGGCTGGGAGACAAAGCTCTCCCTGGA---TC 1079
Oy      1312 CCTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTA 1366
Db      1080 CCTGTGATTTTGAAG--AGATATCTATTTTATTTATTTATTTATTTATTTATTTATTTA 1132

```

RESULT 10
 AC126921
 LOCUS AC126921
 DEFINITION Bos taurus.c1one RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered pieces.

AC126921
 AC126921.1 GI:21724098
 HTG; HTGS_PHASE1; HTGS_DRAFT.
 KEYWORDS
 SOURCE
 ORGANISM
 Bos taurus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovine; Bos.
 1 (bases 1 to 15353)
 Akher, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haglund, P., Hansen, N., Ho, S.-U., Idol, V.R., Karlins, E., Latic, P., Lee-Lin, S.-O., Legaspi, R., Maduro, Q.L., Maduro, V.B., Margulies, E.H., Mastello, C., Maskeri, B., Mastrian, S.D., McLoskey, J.C., McDowell, J., Paguirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Schueler, M.G., Stantirlop, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurgren, C., Vogt, T.L., Walker, M.A., Wetherby, K.D., Wilgins, L., Young, A., Zhang, L.-H., and Green, E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 15353)
 Green, E.D.
 Direct Submission
 Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
 Grovemont Circle, Gailthersburg, MD 20877, USA
 ----- Genome Center
 Center: NIH Intramural Sequencing Center
 Center code: NISC
 Web site: <http://www.nisc.nih.gov>
 Contact: nisc_zoehgrt.nih.gov
 ----- Project Information
 Center project name: ddi
 Center clone name: 045D24
 ----- Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 146066 bases at least Q40
 Consensus quality: 147748 bases at least Q30
 Consensus quality: 148824 bases at least Q20
 Insert size: 15100; agarose-fp
 Insert size: 152453; sum-of-coverage
 Quality coverage: 8.80x in Q20 bases; agarose-fp
 Quality coverage: 8.72x in Q20 bases; sum-of-coverage

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 12 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1
 * 2312: contig of 2312 bp in length
 * 2313
 * 2412: gap of unknown length
 * 2413
 * 5841: contig of 3429 bp in length
 * 5842
 * 5941: gap of unknown length
 * 5942
 * 8435: contig of 2494 bp in length
 * 8436
 * 8535: gap of unknown length
 * 8536
 * 15799: contig of 7264 bp in length
 * 15800
 * 15899: gap of unknown length
 * 15900
 * 25224: contig of 9325 bp in length
 * 25225
 * 32504: gap of unknown length
 * 32525
 * 32504: contig of 7180 bp in length
 * 32505
 * 32604: gap of unknown length
 * 32605
 * 40970: contig of 8366 bp in length
 * 40971
 * 41070: gap of unknown length
 * 41071
 * 56590: contig of 15520 bp in length
 * 56591
 * 56691: gap of unknown length
 * 73769: contig of 17079 bp in length
 * 73770
 * 73869: gap of unknown length
 * 90859: contig of 16990 bp in length

Sequencing vector: Plasmid;
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 112720 bases at least Q40
Consensus quality: 11666 bases at least Q30
Consensus quality: 119165 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
* (see http://www.bgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 1022: contig of 1022 bp in length
1023 1122: gap of unknown length
1123 2394: contig of 1272 bp in length
2395 2494: gap of unknown length
2495 3663: contig of 1169 bp in length
3664 3763: gap of unknown length
3764 4879: contig of 1116 bp in length
4880 4979: gap of unknown length
4980 6787: contig of 1808 bp in length
6788 6887: gap of unknown length
6888 8193: contig of 1305 bp in length
8193 9461: gap of unknown length
9462 9561: contig of 1169 bp in length
9562 11435: gap of unknown length
11436 11535: contig of 1874 bp in length
11536 13535: gap of unknown length
13536 17323: contig of 2000 bp in length
17324 17423: gap of unknown length
17424 20406: contig of 3688 bp in length
20407 20506: gap of unknown length
20507 23445: contig of 2939 bp in length
23446 23560: gap of unknown length
23561 25600: contig of 2055 bp in length
25601 25701: gap of unknown length
25702 28049: contig of 2349 bp in length
28050 28149: gap of unknown length
28150 30474: contig of 2325 bp in length
30475 30574: gap of unknown length
30575 35498: contig of 4924 bp in length
35499 35598: gap of unknown length
35599 39284: contig of 3686 bp in length
39285 39384: gap of unknown length
39385 41970: contig of 2586 bp in length
41971 42070: gap of unknown length
42071 45659: contig of 3589 bp in length
45660 45759: gap of unknown length
45760 50332: contig of 4573 bp in length
50333 50432: gap of unknown length
50433 54695: contig of 4263 bp in length
54696 54795: gap of unknown length
54796 58289: contig of 3494 bp in length
58290 58389: gap of unknown length
58390 63599: contig of 5210 bp in length
63600 63699: gap of unknown length
63700 70305: contig of 6606 bp in length
70306 70405: gap of unknown length
70406 76123: contig of 5718 bp in length
76124 76223: gap of unknown length
76224 84961: contig of 8738 bp in length
84962 85061: gap of unknown length
85062 93614: contig of 8553 bp in length
93615 93714: gap of unknown length
93715 103352: contig of 9638 bp in length
103353 103452: gap of unknown length

FEATURES
source 1. 138792
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-320N23"
BASE COUNT 34391 a 31996 c 31712 g 36193 t 4500 others
ORIGIN
Query Match 18.5%: Score 254.2; DB 2; Length 138792;
Best Local Similarity 66.7%: Pred. No. 1.2e-36;
Matches 517; Conservative 0; Mismatches 188; Indels 70; Gaps 8;
QY 601 CAGGTGCACTTGTATGAGGGGAAAGGCTCTACCTGAAGCTGACTGTGTGATGCT 660
|||||
Db 137627 CAGGTGCACTTGTATGAGGGGAAAGGCTCTACCTGAAGCTGACTGTGTGATGCT 137686
QY 661 GTCTGGCCCTGCGCTGCTGAGAGAAATCTCAGCCACTGCGGCAATTCCTCGGCCC 720
|||||
Db 137687 GTCTGGCCCTGCGCTGCTGAGAGAAATCTCAGCCACTGCGGCAATTCCTCGGCCC 137746
QY 721 CAGTCCGCTCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGG 780
|||||
Db 137747 CAGTCCGCTCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGGCTGCTGAGG 137806
QY 781 ATCCGCAACCTCCCTGGGCGCCATCTCAAGGCTGCGCCCTCTCTCACTTGGAGCTC 840
|||||
Db 137807 ATCCGCAACCTCCCTGGGCGCCATCTCAAGGCTGCGCCCTCTCTCACTTGGAGCTC 137866
QY 841 TTCGAGTTCAGTAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 892
|||||
Db 137867 CACGAAATCAAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 137926
QY 893 GCTCCCTGCAAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 952
|||||
Db 137927 CCAAGACCAACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 137986
QY 953 GCTCCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1012
|||||
Db 137987 GCTCCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138045
QY 1013 ACATTAATACAGTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1072
|||||
Db 138046 ACAGACA-----TATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138100
QY 1073 ACTAGCTCCCAATCCCTGACCTTTGAGGCGCCCAAGTGTGACTGCCCTCGGCCA 1132
|||||
Db 138101 ACGAGCTCCCAATCCCTGACCTTTGAGGCGCCCAAGTGTGACTGCCCTCGGCCA 138138
QY 1133 CAGACCCCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1192
|||||
Db 138139 CCAACCTGCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 138171
QY 1193 CACTTCAAGGCTAAGAGGGCTGACCTGCGGCGAGAGCCAAAGACTGGGCTAG 1252
|||||
Db 138172 TGCATCAGGCTAAGATAGGCTGACCTGCGGCGAGAGCCAAAGACTGGGCTAG 138231
QY 1253 GCCAGGATCCCAATGTGAGGGGCGAGA--AACACACAAGCTGCTGCTGCTGCTGCTGCT 1311
|||||
Db 138232 GCCAGGATCCCAATGTGAGGGGCGAGAAGATGAGACACAAGCTGCTGCTGCTGCTGCT 138287
QY 1312 CCTGTGATTTTAAACAGATATTTATTTATTTATTTATTTATTTATTTATTTATTTATTT 1366
|||||
Db 138288 CCTGTGATTTTGA--AGATACTATTTATTTATTTATTTATTTATTTATTTATTTATTT 138340
RESULT 14

AC130192/c
 LOCUS AC130192 177555 bp DNA linear HTG 08-AUG-2002
 DEFINITION Sus scrofa clone RP44-436K21, WORKING DRAFT SEQUENCE, 10 unordered
 pieces.
 AC130192
 VERSION AC130192.1 GI:22138443
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 REFERENCE
 1 (bases 1 to 177555)
 Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
 Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
 Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
 Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P.,
 Lee-Lin,S.-O., Legaspi,R., Maduro,Q.U., Maduro,V.B.,
 Marquies,E.H., Masiello,C., Maskeri,B., Mestrian,S.D.,
 McLoskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
 Portnoy,M.E., Prasad,A., Schueler,M.G., Stantirlop,S., Thomas,J.W.,
 Thomas,P.J., Touchman,J.W., Tsugeon,C., Vogt,J.L., Walker,M.A.,
 Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
 NISC Comparative Sequencing Initiative
 Unpublished
 2 (bases 1 to 177555)
 Green,E.D.
 Direct Submission
 Submitted (08-AUG-2002) NIH Intramural Sequencing Center, 8717
 Groveom Circle, Gaithersburg, MD 20877, USA

 TITLE NISC
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Center code: NISC
 Center: NIH Intramural Sequencing Center
 Web site: http://www.nisc.nih.gov
 Contact: nisc@coehgri.nih.gov

 Project Information
 Center project name: dbg
 Center clone name: 436K21

 Summary Statistics
 Sequencing vector: plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 164120 bases at least Q40
 Consensus quality: 168085 bases at least Q30
 Consensus quality: 170767 bases at least Q20
 Insert size: 169000; agarose-fp
 Quality coverage: 4.78x in Q20 bases; agarose-fp
 Quality coverage: 4.57x in Q20 bases; sum-of-coverage

 NOTE: This is a 'working draft' sequence. It currently
 consists of 10 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1
 3460: contig of 3460 bp in length
 3461 3560: gap of unknown length
 3561 6869: contig of 3309 bp in length
 6870 6969: gap of unknown length
 6970 13465: contig of 6496 bp in length
 13466 13565: gap of unknown length
 13566 21301: contig of 7736 bp in length
 21302 21401: gap of unknown length
 21402 33719: contig of 12318 bp in length
 33720 33819: gap of unknown length
 33820 45054: contig of 11235 bp in length
 45055 45154: gap of unknown length
 45155 58724: contig of 13570 bp in length
 58725 58824: gap of unknown length
 58825 76670: contig of 17846 bp in length
 76671 76770: gap of unknown length

FEATURES
 source
 1. 177555
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone="RP44-436K21"
 /clone.lib="RP44"
 1. 3460
 /note="assembly_fragment"
 3561. 6869
 /note="assembly_fragment"
 6970. 13465
 /note="assembly_fragment"
 13566. 21301
 /note="assembly_fragment"
 21402. 33719
 /note="assembly_fragment"
 33820. 45054
 /note="assembly_fragment"
 45155. 58724
 /note="assembly_fragment"
 58825. 76670
 /note="assembly_fragment"
 76771. 115165
 /note="assembly_fragment"
 115266. 177555
 /note="assembly_fragment"
 /clone_end:sp6
 /vector_side:right"
 BASE COUNT 41026 a 46997 c 46681 g 41924 t 927 others
 ORIGIN
 Query Match 17.9% Score 246.4; DB 2; Length 177555;
 Best Local Similarity 85.6% Pred. No. 3.2e-35;
 Matches 274; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
 QY 591 CCTGACTGTCTGAGTGTGATGAGGAGGAGGCTGTCTGACCTGAGACTGCT 650
 DB 114241 CCCGCCCCCCCAGGTCGACATTGATGAGGAGGAGGCGCTCTACCTGAGCTGCT 114182
 QY 651 GGTGAGTGTGTGCTGGCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 710
 DB 114181 GGTGAGTGTGTGCTGGCCCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 114122
 QY 711 CCTGGGCCCCCAGTCTCCGCTCTGCGAGGTGTCTGGGCTGTGGCCCTGCGGCGAGGTC 770
 DB 114121 CCTGGGCCCCCAGTCTCCGCTCTGCGAGGTGTCTGGGCTGTGGCCCTGCGGCGAGGTC 114062
 QY 771 CTCCTGCGGAGTCCGACCTCTCCCTGGGCGCATCTCAAGGCTGCCCTTCTCACTCA 830
 DB 114061 CTCCTGCGGAGTCCGACCTCTCCCTGGGCGCATCTCAAGGCTGCCCTTCTCACTCA 114002
 QY 831 CTTGGAGCTCTTCAGTGTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 890
 DB 114001 CTTGGAGCTCTTCAGTGTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 113942
 QY 891 CGGCTCCCTCGACAGCTCT 910
 DB 113941 GGGCTCTCTCTCCCGAGACT 113922
 RESULT 15
 AC126925 161428 bp DNA linear HTG 10-JUL-2002
 LOCUS AC126925/c
 DEFINITION Canis familiaris clone RP81-332E11, WORKING DRAFT SEQUENCE, 15
 unordered pieces.
 ACCESSION AC126925
 VERSION AC126925.1 GI:21724102
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE dog.
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 161428)
AUTHORS Ahlter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granle, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Lario, P., Lee-Jin, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B., Margulies, E.H., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pagulirigan, C., Pearson, R., Portnoy, M.E., Prasad, A., Schueler, M.G., Stancil, P., Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Wehrhby, K.D., Wiggin, L., Young, A., Zhang, L.-H. and Green, E.D.
TITLE NISC Comparative Sequencing Initiative
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 161428)
AUTHORS Green, E.D.
TITLE Direct Submission
JOURNAL Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
COMMENT Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc-zoo@ngri.nih.gov
Project Information
Center project name: CWP
Center clone name: 332E11
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 148712 bases at least Q40
Consensus quality: 151471 bases at least Q40
Consensus quality: 151325 bases at least Q30
Insert size: 15200; agarose-fp
Quality coverage: 5.10x in Q20 bases; sum-of-contigs
Quality coverage: 5.36x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 4734: contig of 4734 bp in length
4735 4834: gap of unknown length
4835 7325: contig of 2491 bp in length
7326 7425: gap of unknown length
7426 12261: contig of 4836 bp in length
12262 12361: gap of unknown length
12362 19031: contig of 6670 bp in length
19032 19131: gap of unknown length
19132 27849: contig of 8718 bp in length
27850 27949: gap of unknown length
27950 31801: contig of 3852 bp in length
31802 31901: gap of unknown length
31902 40781: contig of 8880 bp in length
40782 40881: gap of unknown length
40882 49783: contig of 8902 bp in length
49784 49883: gap of unknown length
49884 59150: contig of 9267 bp in length
59151 59250: gap of unknown length
59251 68679: contig of 9429 bp in length
68680 68780: gap of unknown length
68781 78984: contig of 10205 bp in length
78985 79084: gap of unknown length
79085 93446: contig of 14362 bp in length

FEATURES
source
1. 161428
/organism="Canis familiaris"
/db_xref="taxon:9615"
/clone="RP81-332E11"
/clone="lib="RP81"
1. 4734
/note="assembly-fragment"
misc_feature 4835..7325
/note="assembly-fragment"
misc_feature 7426..12261
/note="assembly-fragment"
misc_feature 12362..19031
/note="assembly-fragment"
misc_feature 19132..27849
/note="assembly-fragment"
misc_feature 27950..31801
/note="assembly-fragment"
misc_feature 31902..40781
/note="assembly-fragment"
misc_feature 40882..49783
/note="assembly-fragment"
misc_feature 49884..59150
/note="assembly-fragment"
misc_feature 59251..68679
/note="assembly-fragment"
misc_feature 68780..78984
/note="assembly-fragment"
misc_feature 79085..93446
/note="assembly-fragment"
misc_feature 93447..102175
/note="assembly-fragment"
/note="assembly-fragment"
clone_end:SP6
vector_side:right
102276..131685
/note="assembly-fragment"
131786..161428
/note="assembly-fragment"

BASE COUNT 39586 a 41863 c 41233 g 37340 t 1406 others
ORIGIN
Query Match 17.7% Score 242.4; DB 2; Length 161428;
Best Local Similarity 64.9%; Pred. No. 1.7e-34;
Matches 504; Conservative 0; Mismatches 201; Indels 71; Gaps 7;
QY 601 CAGGTGACCTTTGATGAGGGAAGGAGCTGTACTGAAGCTGACCTGTGGATGATGT 660
|||||
DB 42487 CAGGTGACCTTTGATGAGGGAAGGAGCTGTACTGAAGCTGACCTGTGGATGATGT 42428
QY 661 GTGCTGACCTTGTGAGGAGGAATCTCAAGCAGCTGCGGCACTTCCCTGGGCCC 720
|||||
DB 42427 GCCCTGACCTTGTGAGGAGGAATCTCAAGCAGCTGCGGCACTTCCCTGGGCCC 42368
QY 721 CAGCTGCGCTTGTGAGGAGGAATCTCAAGCAGCTGCGGCACTTCCCTGGGCCC 780
|||||
DB 42367 CAGCTGCGCTTGTGAGGAGGAATCTCAAGCAGCTGCGGCACTTCCCTGGGCCC 42308
QY 781 ATCCGACCTTGTGAGGAGGAATCTCAAGCAGCTGCGGCACTTCCCTGGGCCC 840
|||||
DB 42307 ATCCGACCTTGTGAGGAGGAATCTCAAGCAGCTGCGGCACTTCCCTGGGCCC 42248
QY 841 TTCAGGTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
|||||
DB 42247 TTCAGGTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 42212
QY 901 CGACAGCTTCTGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 960
|||||

D	42211	CTGCAGCTCCGCCACCCCGCTCCCTGACCTGCGTCCCTCTGACGAGCGCTGCTGGACCTTG	42152
O	961	CTGCGCCCTCCCTCTAGAGGCTGCTGGGCTCTTTCACGTGTTTTCATCCACATTAAT	1020
D	42151	TTACCTCTGTTTCCGGCCACATACATTAATCTCTTCTCCCTCTGACACCCCGCATCT	42096
O	1021	ACAGATATCCACACTTATATCTTAACAATCCCGCCACGGCCACTCTCAACTCATCAGGTC	1080
D	42095	-----CCACCTCACTGGCCCCCTGAACCCCTTGCTTTTGAAGCCCCCATTTATCTC	42045
O	1081	CCCAATCCCTGACCCCTTTGAGGCCCCCAGATCTCGACTCCCCCTGGCCACAGACCC	1140
D	42044	CTGACTTCTCACTCTTGCCCAACCCCGCCACCCCGCTCCGCCCAACACAGC----	41989
O	1141	CAGGCATTGTGTTCACGTCTACTCTGTGGGCGAAGATGGGTGCACA---GACCCCACTT	1197
D	41988	-----CAAAGCACACTGTGTCTATTCTTGGGTAGAAATGGGTCCCTTAACCTCCGCCACTT	41934
O	1198	CAGCAGCTAAGAGGGGCTGACCTGGCGGACAGAAAGCCAAAGACACTGGGCTTAGGCAG	1257
D	41933	CAGGCACATATACAGGGCTAGACGCCAGGGCGACTGGGACAGC-----GAGGCCGG	41884
O	1258	GAGTTCCTCCAAATGTGAGGGGCGAGAAACAGCAAGCTCTCCCTTGAAATTTCCCTGTG	1317
D	41883	GAGTTCCTCCATCGCCAGAGGGCTGAG--AACAGAGTAGAGTCC-----TTCCCTGTG	41836
O	1318	GATTTTAAACAGATATTTATTTATATATATTTGTGCAAAATGTGTATATAATG	1373
D	41835	GATTTTAAACAGATATTTATTTATATATATTTGTGCAAGATGTCAATTAATGTG	41780

Search completed: March 31, 2003, 11:49:23
Job time : 6804.73 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 05:28:59 ; Search time 42.8782 Seconds
(without alignments)
1081.216 Million cell updates/sec

Title: US-09-245-198a-2
Perfect score: 1162
Sequence: 1 VLISGLALACIGLLVVS.....PMAHLKAAPFLTYGELFQVH 225

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	107.5	9.3	325	5 Q9V5G2	Q9V5G2 drosophila
2	93.5	8.0	557	16 Q8XQX3	Q8XQX3 ralsstonia s
3	90.5	7.8	210	16 Q9A926	Q9A926 caulobacter
4	90.5	7.8	421	16 Q9HWW2	Q9HWW2 pseudomonas
5	89.5	7.7	805	5 Q9VPD4	Q9VPD4 drosophila
6	89	7.7	287	13 Q90W79	Q90W79 gallus gall
7	89	7.7	352	12 Q89341	Q89341 hendra viru
8	88	7.6	224	5 Q9V762	Q9V762 drosophila
9	88	7.6	353	12 Q66760	Q66760 equine morb
10	88	7.6	532	4 Q16727	Q16727 homo sapien
11	88	7.6	1663	4 Q9U001	Q9U001 homo sapien
12	88	7.6	3261	4 Q9Y556	Q9Y556 homo sapien
13	88	7.6	3664	4 Q96T58	Q96T58 homo sapien
14	87.5	7.5	522	10 Q9FTN7	Q9FTN7 oryza sativ
15	87.5	7.5	670	16 Q9AA15	Q9AA15 caulobacter
16	86.5	7.4	340	16 Q9HUR8	Q9HUR8 pseudomonas

17	86.5	7.4	504	16 Q92KA4	Q92KA4 rhizobium m
18	86	7.4	260	10 Q8S2M9	Q8S2M9 oryza sativ
19	85.5	7.4	331	10 Q942P9	Q942P9 oryza sativ
20	84.5	7.3	422	16 Q9RKB0	Q9RKB0 streptomyce
21	84.5	7.3	430	2 Q9RE01	Q9RE01 streptomyce
22	84.5	7.3	1079	13 Q8UVR4	Q8UVR4 xenopus lae
23	84.5	7.3	1118	16 Q98E34	Q98E34 rhizobium l
24	84	7.2	2962	5 Q93326	Q93326 caenorhabdi
25	83.5	7.2	394	16 Q92V66	Q92V66 rhizobium m
26	83	7.1	467	16 Q9S2Y4	Q9S2Y4 streptomyce
27	83	7.1	2779	5 Q9W4N7	Q9W4N7 drosophila
28	82.5	7.1	174	16 Q9CXX1	Q9CXX1 pasteurella
29	82.5	7.1	314	5 Q8WP77	Q8WP77 theileria a
30	82.5	7.1	549	16 Q8RC38	Q8RC38 thermoaer
31	82	7.1	619	5 Q8SU99	Q8SU99 encephalit
32	82	7.1	654	16 P74664	P74664 synechocyst
33	81.5	7.0	999	11 Q9JXR6	Q9JXR6 mus musculu
34	81.5	7.0	1665	11 Q8V111	Q8V111 mus musculu
35	81	7.0	258	4 Q9NUD0	Q9NUD0 homo sapien
36	81	7.0	262	11 Q9CVL1	Q9CVL1 mus musculu
37	80.5	6.9	274	10 Q949L6	Q949L6 beta vulgar
38	80.5	6.9	324	11 Q9JUS4	Q9JUS4 rattus norv
39	80.5	6.9	788	4 Q9P2E5	Q9P2E5 homo sapien
40	80.5	6.9	1438	16 Q8S739	Q8S739 pseudomonas
41	80	6.9	319	16 Q8U8B6	Q8U8B6 agrobacteri
42	80	6.9	403	6 Q9BGL8	Q9BGL8 canis famli
43	80	6.9	589	5 Q8T0B2	Q8T0B2 drosophila
44	80	6.9	793	5 Q9Y108	Q9Y108 drosophila
45	80	6.9	1522	10 Q8S4C3	Q8S4C3 oryza sativ

ALIGNMENTS

RESULT 1

Q9V5G2	PRELIMINARY;	PRT;	325 AA.
ID Q9V5G2	AC Q9V5G2	DT 01-MAY-2000 (TREMBLrel. 13, Created)	DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)	DE CG12919 protein.	OS CG12919.	OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;	OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	OC Ephydroidea; Drosophilidae; Drosophila.	OX NCBI_TaxID=7227;
OX [1]	RN	RP	SEQUENCE FROM N.A.
RC STRAIN=BERKELEY;	RX MEDLINE=20196006; PubMed=10731132;	RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,	RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abiril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,	RA Balles R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,	RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Dudin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,	RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,	RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin K.A., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,	RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,	

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mlshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzo M., Piltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spadling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.H., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gidbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AEO03831; AAF5848.1; -
 DR FLYbase; FBgn0033483; CG12919.
 DR InterPro; IPR000478; TNF_family.
 DR SMART; SM00207; TNF_1.
 DR PROSITE; PS00251; TNF_1; 1.
 DR PROSITE; PS50049; TNF_2; 1.
 DR PROSITE; PS50049; TNF_2; 1.
 SO SEQUENCE 325 AA; 36862 MW; 6E5CCB6964F1A3A CRC64;

Query Match 9.3%; Score 107.5; DB 5; Length 325;
 Best Local Similarity 23.1%; Pred. No. 0.046;
 Matches 48; Conservative 34; Mismatches 91; Indels 35; Gaps 9;

OY 29 QEPQDELTDREPELNPQTEESQDVVPLEQLVPRRS--APGRKAPRAIAAH 86
 DB 141 QEKSNATSKESPAPLHRRMRSH-----RHLLVRKGESLLARSSESRP---AAH 191
 OY 87 YEVPAPRGQGAQGVDTGVSQMEETKINSSPLRYDROIGEFVIRAGLYLVCOVHP 146
 DB 192 FHLSRRRHGSM-GYHGDMYIGNDNERNSYG-HFQTRGVLTVTGTLGYVAQICYN 249
 OY 147 EGRAVYLLKDLVNGVLA-----LRCLFEFSATTAASPEPQIRLCQVSGLLPLRPGSS 199
 DB 250 NSHD-----QNGFTVFQGDTPFLQCLN---TVPTNPMHKVHTCHTSGLLHLERNER 297
 OY 200 LRITL---PWHLKAPFLTYFGLFOV 224
 DB 298 IHLKDINHNRNAVLRGNNSYFGLFV 325

RESULT 2

OBXOX3 PRELIMINARY; PRT; 557 AA.

AC OBXOX3;
 DT 01-MAR-2002 (TEMBLrel. 20, Created)
 DT 01-MAR-2002 (TEMBLrel. 20, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE Probable activation/secretion signal peptide protein.
 GN RSP1095 OR RS02601.
 OS Ralstonia solanacearum (Pseudomonas solanacearum).
 OG Plasmid megaplasmid.
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-GM1100;
 RX MEDLINE=21681879; Pubmed=11823852;
 RA Sallanoubet M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlet M., Billault A., Brotier P., Camus J.C., Catolico L.,
 RA Chandler M., Choisme N., Claudel-Renard C., Cunnac S.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Siguer P., Thebaud P., Whalen M., Wincker P., Levy M.,
 RA Weisenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*.";
 RL Nature 415:497-502(2002).
 DR EMBL; AL646082; CAD18246.1; -.

KW Plasmid; Complete proteome.
 SO SEQUENCE 557 AA; 60786 MW; 8B974C147D710649 CRC64;

Query Match 8.0%; Score 93.5; DB 16; Length 557;
 Best Local Similarity 26.6%; Pred. No. 2;
 Matches 62; Conservative 25; Mismatches 77; Indels 69; Gaps 14;

OY 2 LSLGLALACGLLLVVSLSGMATLS-AQEPQDELTAEDRREPPE--LNPQTEESQDVV 58
 DB 10 LALGVALA-----GFLSSMAQSSPACGNP-IDTLPRVDTSPRPCKIHVQGRN--- 58
 OY 59 PTEBQL---VPRRSAPKGRKAPRAIAAHYVHPRGQGAQGVDTGVSQMEETKI 114
 DB 59 PALENLTATLTPESKQIEGVKALPPEELIAHFA--PMAGHDVTVAQLQAA--EVTKL 114
 OY 115 --NSSSPLRYDROIGEFVIRAGLYLVCOVHDEKAVYLLKDLVNGVLAIRCLEERS 172
 DB 115 YADRGYPL-----SAFPAQTF-----EGGVARI---TVEGVIAKRIE--- 152
 OY 173 ATAAASPFP---QLRLC-----QVSGLLPLRPGSSLRITLP 206
 DB 153 ----GKRGPLEARLRATSKHMDERPLRRETFERYTGVALQGVQITATVP 201

RESULT 3

OYA926 PRELIMINARY; PRT; 210 AA.

AC OYA926;
 DT 01-JUN-2001 (TEMBLrel. 17, Created)
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
 DT 01-MAR-2002 (TEMBLrel. 20, Last annotation update)
 DE Hypothetical protein CC1168.
 GN CC1168.
 OS Caulobacter crescentus.
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
 OC Caulobacter.
 OX NCBI_TaxID=155892;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ATCC 19089 / CB15;
 RX MEDLINE=21173698; Pubmed=11259647;
 RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadok N.D., Ely B.,
 RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
 RA Kolonay J.F., Smit J., Craven M.B., Khouli H., Shetty J., Berry K.,
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
 RT "Complete genome sequence of *Caulobacter crescentus*.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
 DR EMBL; AE005794; AAK23152.1; -.
 DR TIGR; CC1168; -.
 KW Hypothetical protein; Complete proteome.
 SO SEQUENCE 210 AA; 22344 MW; 8B9830ADF87F45C CRC64;

Query Match 7.8%; Score 90.5; DB 16; Length 210;
 Best Local Similarity 26.2%; Pred. No. 1.1;
 Matches 33; Conservative 19; Mismatches 67; Indels 7; Gaps 2;

OY 2 LSLGLALACGLLLVVSLSGMATLS-AQEPQDELTAEDRREPPELNPQTEESQDVVPL 61
 DB 32 LRLGLAVFAAGVATVIVQAMRSCLASSKLOTAAAPLVLDPKPR--TGVLKDSRPFL 88
 OY 62 EQLVPRPSAPKGRKAPRAIAAHYVHPRGQGAQGVDTGVSQMEETKINS 117
 DB 89 ITRERERDRAKDONIVRLTAPLVKRGYGEPPNSQATKASGVYREAEVTLTLDIVKATISA 148
 OY 118 SPLRYD 123
 DB 149 EGDFD 154

RESULT 4


```

Q9H0W2          PRELIMINARY;      PRT:      421 AA.
ID  Q9H0W2
AC  Q9H0W2:
DT  01-MAR-2001 (TReMBLrel. 16, Created)
DT  01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT  01-OCT-2001 (TReMBLrel. 18, Last annotation update)
DE  Hypothetical protein PA4851.
GN  PA4851.
OS  Pseudomonas aeruginosa.
OC  Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OX  Pseudomonas
NCBI_TaxID=287;
RN  NCBI_TaxID=287;
RP  [1]
RC  SEQUENCE FROM N.A.
RX  STRAIN=ATCC 15692 / PA01;
MEDLINE=20437337; PubMed=10984043;
RA  Steyer C.K., Pham X.-Q.T., Ervin A.L., Mizoguchi S.D., Warren P.,
RA  Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA  Garber R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA  Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lam R.M.,
RA  Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA  Reizer J., Slier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT  "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT  opportunistic pathogen.";
RL  Nature 406:959-964(2000).
DR  EMBL; AE004898; AAC08236.1;
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 421 AA; 46442 MW; 68C5C17099953C3E CRC64;

Query Match          7.8%; Score 90.5; DB 16; Length 421;
Best Local Similarity 22.6%; Pred. No. 2.7;
Matches 53; Conservative 27; Mismatches 76; Indels 79; Gaps 10;

OY  29 QEPSEELTAADR-----REPELNPQTESQDVVPL-----EQLV 65
DB  163 QEPSEGLAEERSEFVPLRAEDPEPLRLSDDDDEPSGKRRLGLSALDEDESSA 222
OY  66 RPRSPASGRKARPRRAIAAHYVHPRPGODGAGVGTSGMEETKINSSPLRYDRO 125
DB  223 EEGSVAPLGSARKRDEA-----PVESLDQLRDEPLQLAWMKPR-----KQWRR 266
OY  126 IGEFTVI---RAGLYYYLYCQVHFE-----GKAVYLIKLDL--LVNGV 162
DB  267 LGWLILILALGGLAAQYIAVHFDELARQDARVPMFAOLCEPICGCTLSKVDVEQIRSSN 326
OY  163 LALRCLREESA-----TASSPGP--QLRLCQVSGLL-----PLRGSSSL 200
DB  327 LVKRSHPFSGALVVDATITYNRASFSPPLLELRFADLNGHLIANRRPKPGEYL 381

RESULT 5
O9YFD4          PRELIMINARY;      PRT:      805 AA.
ID  O9YFD4
AC  O9YFD4:
DT  01-MAY-2000 (TReMBLrel. 13, Created)
DT  01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT  01-JUN-2000 (TReMBLrel. 14, Last annotation update)
DE  CG18442 protein.
GN  CG18442.
OS  Drosophila melanogaster (Fruit fly).
OC  Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC  Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC  Ephydroidea; Drosophilidae; Drosophila.
OX  NCBI_TaxID=7227;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=BERKELEY;
RX  MEDLINE=20196006; PubMed=10731132;
RA  Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA  Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA  George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA  Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA  Brandon R.C., Rogers Y.-H.C., Blazek R.G., Champe M., Pfeiffer B.D.,

```

```

RA  Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA  Abril J.F., Abgayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA  Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA  Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolintsov S.,
RA  Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
RA  Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA  Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA  de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA  Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA  Durbin K.J., Evangelista C.C., Ferraz C., Ferlita S., Fleischmann W.,
RA  Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA  Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA  Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA  Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA  Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA  Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Liap D., Lai Z.,
RA  Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA  Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA  Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA  Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA  Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclob J.M.,
RA  Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA  Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA  Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
RA  Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA  Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA  Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA  Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA  Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA  Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA  Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT  "The genome sequence of Drosophila melanogaster.";
RL  Science 287:2185-2195(2000).
DR  EMBL; AE003707; AA055126.1;
DR  FlyBase; FBgn0038287; CG18442.
SQ  SEQUENCE 805 AA; 86321 MW; FDF9FEF14E69957 CRC64;

Query Match          7.7%; Score 89.5; DB 5; Length 805;
Best Local Similarity 21.7%; Pred. No. 7.7;
Matches 49; Conservative 27; Mismatches 55; Indels 95; Gaps 10;

OY  29 QEPSEELTAADRREPELNPQTESQDVVPLLEQVAPRRSAPRKARPR--ATAAH 86
DB  482 EQPGQEEI-----QPPRPTPTTEQ-----LSPPARPKSAELLQYSPKKQVRIAS 529
OY  87 YEYHRR-----PGQDGAQAGVDGYTSGHEETKI-----N 115
DB  530 PVMQPERRELCPPLPPRGSPITLDGSSQSPTNVAVSGPKKPLPIACRPPRSGVNSPN 589
OY  116 SSS-----PLRYDRI-----GFTVIRAGLYLYCOV 143
DB  590 SSSPSAPPAPHTSPPIPTVRLPHLNQANGTLPLPKPQOLHGKFLKNG----- 641
OY  144 HFDEGKAVYLIKLDLVNGVL-----ALRCLSEFSATAASPGPOL 183
DB  642 HLIDDEALLAKTDVAMSGLLIKLDVQAQC-----SVAQAAGGTFI 683

RESULT 6
O90WT9          PRELIMINARY;      PRT:      287 AA.
ID  O90WT9
AC  O90WT9:
DT  01-DEC-2001 (TReMBLrel. 19, Created)
DT  01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT  01-MAR-2002 (TReMBLrel. 20, Last annotation update)
DE  TNF-related apoptosis inducing ligand.
OS  Gallus gallus (Chicken).
OC  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC  Gallus; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OX  NCBI_TaxID=9031;
RN  [1]
RP  SEQUENCE FROM N.A.

```

RA Brdgham J.T., Johnson A.L.;
 RT "TNF-related apoptosis inducing ligand (TRAIL) expression in the hen
 ovary.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY057941; AAL23702.1; -;
 DR InterPro: IPR003636; TNF_abC.
 DR pfam: PF00229; TNF; 1.
 DR ProDom: PD002012; TNF_abC; 1.
 DR PROSITE: PS50049; TNF-2; 1.
 SQ SEQUENCE 287 AA; 32092 MW; DB06ELC95087B108 CRC64;

Query Match 7.7%; Score 89; DB 13; Length 287;
 Best Local Similarity 23.4%; Pred. No. 2.3;
 Matches 40; Conservative 24; Mismatches 65; Indels 42; Gaps 7;

OY 84 AAHYEHPR-PGDDGAGAGD-----TVSGWEETKINS-SPLRYDQIGEPYIRA 134
 DB 127 SAHLIRPQNPADGSSRRFGNLSQSCRAHATWEDSTHSHLQNTY--RDGRLRVNQA 184
 OY 135 GLLYLQVNF-----EGKAVYLKDLVNGVIALRCLFEFSATA 175
 DB 185 GKYYVSQIFRYSRQAGARVSPQVOCINMKTYSQPIILLKCV-----GTK 234
 OY 176 ASGPGQ--LRICVSGLELRPGSSLRIRTLPLMAHLKAAPLTYGGLFOV 224
 DB 235 CMAPEAEVGLHALYOGLELKADELIVSVSLAIDYSDAASVYGAFRL 285

RESULT 7

ID 089341 PRELIMINARY; PRT; 352 AA.

AC 089341;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Matrix.

OS Hendra virus.
 OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
 OC Paramyxoviridae; Paramyxovirinae; Morbilliviruses.

OX NCBI_TaxID=63330;
 RN RN
 RP SEQUENCE FROM N.A.
 RA MEDLINE=9058172;
 RA Yu M., Hansson E., Langedijk J.P., Eaton B.T., Wang L.F.;
 RT "The attachment protein of Hendra virus has high structural similarity
 RT but limited primary sequence homology compared with viruses in the
 RT genus Paramyxovirus.";
 RL Virology 251:227-233(1998).

RL (2)
 RN RN
 RP SEQUENCE FROM N.A.
 RA MEDLINE=20481636; PubMed=11024125;
 RA Wang L.F., Yu M., Hansson E., Pritchard L.I., Shnell B.,
 RA Michalski W.P., Eaton B.T.;
 RT "The exceptionally large genome of Hendra virus: support for creation
 RT of a new genus within the family Paramyxoviridae.";
 RL J. Virol. 74:9972-9979(2000).

RN RN
 RP SEQUENCE FROM N.A.
 RA Wang L.F., Yu M., Pritchard L.I., Hansson E., Eaton B.T.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF017149; AAC83191.1; -;
 DR InterPro: IPR000982; Matrix.
 DR pfam: PF00661; Matrix; 1.
 DR ProDom: PD000741; Matrix; 1.
 SQ SEQUENCE 352 AA; 39793 MW; 79E238DE496828D5 CRC64;

Query Match 7.7%; Score 89; DB 12; Length 352;
 Best Local Similarity 28.1%; Pred. No. 3;
 Matches 39; Conservative 20; Mismatches 50; Indels 30; Gaps 8;

OY 78 RPRRAIAHYEHPRPGDAGAGVDTVSGWEETKINS-SPLRYDQIGEPYIRA 137

DB 197 RRRNNAIFNLVLYLKIDADLAKAGISFDR-DGTVAS-----FMLHIGN-VRRAGYV 249
 OY 138 YL--YQVHFDECKANY-----LKLIDLNVNGVIALRCLFEFSATRASSPGPOLRCOV 188
 DB 250 YSVEYCKRRKIDRKLQFSGISGGLSLHTKINQISKRLPAQW-----GLQKMLC-- 299
 OY 189 SGLPLRPGSSLRIRTLPLPW 207
 DB 300 FSLMDINPW---LNRLLTW 314

RESULT 8

ID 09V762 PRELIMINARY; PRT; 224 AA.

AC 09V762;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
 DE CG10205 protein.

GN CG10205.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.

OX NCBI_TaxID=7227;
 RN RN
 RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Ceiniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agdayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu U., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhattacharya D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,
 RA Burks K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fouts R.C., Gabrielian A.E., Gary N.S., Gilbert W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jajalil M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Laske P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mikhlin N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusslein B., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein D.A., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003813; AAF58200.1; -;
 DR Flybase: FBgn0033970; CG10205.
 SQ SEQUENCE 224 AA; 25025 MW; 780B9EF7BBAFF156 CRC64;

Query Match 7.6%; Score 88; DB 5; Length 224;

Best Local Similarity 22.6%; Pred. No.2.1;
Matches 44; Conservative 27; Mismatches 52; Indels 72; Gaps 10;

OY 10 CUGLLVYVSLGSAWATLSAQPSOEELTAED-----KREPELLNQ--TEESQ 55
DB 13 CIGAL-----TWASLAIEESTDIE---EDVISTANOTQYIRVHPDLPKPKDTTGSQ 62
OY 56 DVVPLE-OLVPRRS-----APGGRKAPRRA-----TAHVEHPRGQDGA 98
DB 63 MNSPTQVOTLKPSTAFSSQROQYVDSKOMRRRPPALGCVAA-----SGEDAT 115
OY 99 QAGVGTWVGMEETKINSSPLRYD-----RQIGFTVIRAG----- 135
DB 116 QPALYELKAFPKOKLKQKQKQNLIDARAGELASAGQVQLKQVEVGGVLPVQMTPEPY 175
OY 136 -LYLYICOVHFEDEK 149
DB 176 PIYVYVSKTNGRFGK 190

RESULT 9

ID 066760 PRELIMINARY; PRT; 353 AA.
AC 066760;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Matrix protein.
OS Equine morbillivirus.
OC Viruses; ssRNA negative-strand viruses; Mononegavirales;
OC Paramyxoviridae; Paramyxovirinae; Morbillivirus.
OX NCBI_TaxID=45337;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96419932; PubMed=8822631;
RA Gould A.R.;
RT "Comparison of the deduced matrix and fusion protein sequences of
RT equine morbillivirus with cognate genes of the Paramyxoviridae.";
RL Virus Res. 43:17-31(1996).
DR EMBL; U49404; AAB39504.1; -
DR InterPro; IPR000982; Matrix.
DR Pfam; PF00661; Matrix; 1.
DR ProDom; PD000741; Matrix; 1.
SQ SEQUENCE 353 AA; 40142 MW; EBB67A1AB0FB2AD CRC64;

Query Match 7.6%; Score 88; DB 12; Length 353;
Best Local Similarity 28.1%; Pred. No.3.7;
Matches 39; Conservative 20; Mismatches 50; Indels 30; Gaps 8;

OY 78 RPRRAIAAHVEYHPRPGDQAGVDTVSGWETKINSSPLRYDROIGFTVIRAGLY 137
DB 198 RNNNAIAFRLVLYKIDADLAKAGTGSFDR-DGTRKVAS-----FMHLGNF-VRRAGKY 250
OY 138 YL--YCOVHFDEGKAVY-----LKLDLVNGVLAALRCLEEFSAATASSPQQLRLCOV 188
DB 251 YSEVECKRRIDMKLOFSIGSIGSLHIKINGVISKRLFAQM-----GFOKNLC-- 300
OY 189 SGLPLRPGSSSLRIRLTPW 207
DB 301 PSLMDINPW---LNRLLTW 315

RESULT 10

ID 016727 PRELIMINARY; PRT; 532 AA.
AC 016727;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Placental-like alkaline phosphatase precursor (EC 3.1.3.1).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON;
RX MEDLINE=90124311; PubMed=2297757;
RA Gum J.R., Hicks J.W., Sack T.L., Kim Y.S.;
RT "Molecular Cloning of Complementary DNAs Encoding Alkaline Phosphatase
RT in Human Colon Cancer Cells";
RL Cancer Res. 50:1085-1091(1990).
DR EMBL; X53279; CA37374.1; -
DR HSSP; P00634; IADC.
DR InterPro; IPR001952; Alk_phosphatase.
DR Pfam; PF00245; alk_phosphatase; 1.
DR PRINTS; PR00113; ALKPHPTASE.
DR ProDom; PD001868; Alk_phosphatase; 1.
DR SMART; SM00098; alkppc; 1.
DR PROSITE; PS00123; ALKALINE_PHOSPHATASE; 1.
KW Signal.
FT SIGNAL 1
FT CHAIN 20 POTENTIAL.
SQ SEQUENCE 532 AA; 57399 MW; 354B345CB19A0BC4 CRC64;

Query Match 7.6%; Score 88; DB 4; Length 532;
Best Local Similarity 28.6%; Pred. No.6.3;
Matches 44; Conservative 16; Mismatches 44; Indels 50; Gaps 8;

OY 76 KAPRRRAIAAHVEY--HPRG---QDGAQAGVDGTWVGMEETKINSSPLRYDROIGEF 129
DB 400 KARDRKA---YTVLLYNGPGYVLKDGARPDTESESGSPRYQSAVPLDGETHAGED 455
OY 130 TVIRAGLYLYCOVHFDEGKAVLYLKLDLVNGV-----LALRCLEEPSA----- 173
DB 456 VAV-----FANGPQAH-----LVHGVQGTFTAHYMAFACLEPTACDLAPS 498
OY 174 ---TAASSPQQLRLCOVSGLLPLRPGSSLRIT 204
DB 499 AGTTDAHFGPSV---VPALLPLLAGTLLILGT 528

RESULT 11

ID 090001 PRELIMINARY; PRT; 1663 AA.
AC 090001;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE KIA00929 protein (Fragment).
CN KIAA0929.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RX MEDLINE=99246063; PubMed=10231032;
RA Nagase T., Ishikawa K., Suyama M., Kikuno R., Hirose M.,
RA Miyajima N., Tanaka A., Kotani H., Nomura N., Ohara O.;
RT "Prediction of the coding sequences of unidentified human genes. XIII.
RT The complete sequences of 100 new cDNA clones from brain which code
RT for large proteins in vitro.";
RL DNA Res. 6:63-70(1999).
DR EMBL; AB023146; BAA76773.1; -
FT NON_TER 1
SQ SEQUENCE 1663 AA; 173871 MW; EC94D17086A8531E CRC64;

Query Match 7.6%; Score 88; DB 4; Length 1663;
Best Local Similarity 23.1%; Pred. No.27;
Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;

OY 33 QBELTAEDRREPELLNPQTESQ-----DVVPLEOLVPRRSAPK----- 73
DB 1400 QRLPAGPANRPEPHPTQVOROAETGPTSPSPVSVSMKPLD-LVSLPTQAPKQPLFV 1458

```

OY 74 ---GRKARPRRAIAAHYEVHPRPGDGAAGVDTGTVSGMEETKINSSEPLRYDRQIGFT 130
DB 1459 PTTSGSTPGGLVLPHTTEFPAPKQDS-----PHLTSGRPVDMVQLKKY 1505
OY 131 VIRAGLYYL-----YCOVHFDESKAVYLLKIDLNVGVLCLEEFSAATAASPGQLRL 185
DB 1506 IVMQGLLALKNDTAAYQLHVFSG-----NNVLAHRSLL-----PLSGGPPRLRI 1548
OY 186 CQ 187
DB 1549 AQ 1550

RESULT 12
OY556 PRELIMINARY: PRT: 3261 AA.
AC 09Y556;
DB 01-NOV-1999 (TREMBLrel. 12, Created)
DB 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DB 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 357.0 kDa protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCB1_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA Rhodes S., Huckle E.;
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL096858; CAB51072.1;
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 2.
DR SMART: SM00360; rrm; 2.
DR PROSITE: PS50102; RRM; 2.
KW Hypothetical protein.
SQ SEQUENCE 3261 AA: 357037 MW: F8B6A645DD9B6BC CRC64;

Query Match 7.6%; Score 88; DB 4; Length 3261;
Best Local Similarity 23.1%; Pred. No. 65;
Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;

OY 33 QEELTADREPELNPQTEESQ-----DVYPLEQLVRRRSAPK----- 73
DB 2998 QPRLPGPANRPPEPHTVQVRAQETGPTSPSPSVSKMPLD-PVSLPTQTAPKQPLEV 3056
OY 74 ---GRKARPRRAIAAHYEVHPRPGDGAAGVDTGTVSGMEETKINSSEPLRYDRQIGFT 130
DB 3057 PTTSGSTPGGLVLPHTTEFPAPKQDS-----PHLTSGRPVDMVQLKKY 3103
OY 131 VIRAGLYYL-----YCOVHFDESKAVYLLKIDLNVGVLCLEEFSAATAASPGQLRL 185
DB 3104 IVMQGLLALKNDTAAYQLHVFSG-----NNVLAHRSLL-----PLSGGPPRLRI 3146
OY 186 CQ 187
DB 3147 AQ 3148

RESULT 13
OY556 PRELIMINARY: PRT: 3664 AA.
AC 09Y556;
DB 01-DEC-2001 (TREMBLrel. 19, Created)
DB 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DB 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Nuclear receptor transcription cofactor.
GN SHARP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCB1_TaxID=9606;
RN 11

```

```

RP SEQUENCE FROM N.A.
RA Shi Y., Downs M., Xie W., Kao H.-Y., Ordentlich P., Tsai C.-C.,
RA Hon M., Evans R.M.;
RT "SHARP, an inducible cofactor that integrates nuclear receptor
RT repression and activation."
RL Genes Dev. 0:0-0(2001).
DR EMBL; AF356524; AK52750.1;
DR InterPro: IPR000504; RNA_rec_mot.
DR Pfam: PF00076; rrm; 4.
DR PROSITE: PS50102; RRM; 4.
KW Receptor.
SQ SEQUENCE 3664 AA: 402243 MW: 5228C58533E5B27B CRC64;

Query Match 7.6%; Score 88; DB 4; Length 3664;
Best Local Similarity 23.1%; Pred. No. 75;
Matches 42; Conservative 20; Mismatches 62; Indels 58; Gaps 7;

OY 33 QEELTADREPELNPQTEESQ-----DVYPLEQLVRRRSAPK----- 73
DB 3401 QPRLPGPANRPPEPHTVQVRAQETGPTSPSPSVSKMPLD-PVSLPTQTAPKQPLEV 3459
OY 74 ---GRKARPRRAIAAHYEVHPRPGDGAAGVDTGTVSGMEETKINSSEPLRYDRQIGFT 130
DB 3460 PTTSGSTPGGLVLPHTTEFPAPKQDS-----PHLTSGRPVDMVQLKKY 3506
OY 131 VIRAGLYYL-----YCOVHFDESKAVYLLKIDLNVGVLCLEEFSAATAASPGQLRL 185
DB 3507 IVMQGLLALKNDTAAYQLHVFSG-----NNVLAHRSLL-----PLSGGPPRLRI 3549
OY 186 CQ 187
DB 3550 AQ 3551

RESULT 14
OY556 PRELIMINARY: PRT: 522 AA.
AC 09Y556;
DB 01-MAR-2001 (TREMBLrel. 16, Created)
DB 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DB 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative Rer1A protein (AtRER1A).
GN P0005A05.12.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Cerealia; Oryzae; Oryza.
ON NCB1_TaxID=4530;
RN 11
RP SEQUENCE FROM N.A.
RA STRAIN-CV, NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone: P0005A05.12."
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP002863; BAB16908.1;
DR InterPro: IPR004932; Rer1.
DR Pfam: PF03248; Rer1; 1.
SQ SEQUENCE 522 AA: 58171 MW: AE43E294DFE8BE7F CRC64;

Query Match 7.5%; Score 87.5; DB 10; Length 522;
Best Local Similarity 30.3%; Pred. No. 6.9;
Matches 57; Conservative 21; Mismatches 57; Indels 53; Gaps 15;

OY 70 SAKGKARPRRAIAAHYEVHPRPGDGAAGVDTGTVSGMEETKINSSEPLRYDRQIGFT 113
DB 201 SSSKGRGEPARVVA-----SPSGFRPGFSMDAADSSTAAGAAAAMKMTDASRAFY 256
OY 114 INSSPLRYDRQIGFTYIRAGLYYLCOVHFDESKAVYLLKIDLNVGVLCLEEFSAATAASPGQLRL 166
DB 257 LDRSTHATGRWIGTILAV--AAIYAL--RYVLQGVYIYTGIGITL--LNLIL--GLLSLM 310
OY 167 CLEEFSAATAASPGQLRLCOVSGLLPLRPGSSLR--IRTLP-----WAHLKA---APPL 216

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 01:56:48 ; Search time 1488.36 seconds

(without alignments)
2448.323 Million cell updates/sec

Title: US-09-245-198a-2
Perfect score: 1162

Sequence: 1 VLSGLALACGLLVVSL.....PMAHLKAPFLYFGFQVH 225

Scoring table:

BIOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 segs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/sgn2.1/USPRO/spool/US09245198/runat_24032003.163554.25956/app.query.fasta.1.846
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blissum62 -FRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09245198 @CGN.1.1.2463 @runat.24032003.163554.25956 -MCPU=6 -ICPU=3
-NO_XIPXY -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -MATT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST : *
1: em_estba : *
2: em_esthum : *
3: em_estlin : *
4: em_estlmu : *
5: em_estlov : *
6: em_estlpl : *
7: em_estro : *
8: em_hic : *
9: gb_est1 : *
10: gb_est2 : *
11: gb_hic : *
12: gb_est3 : *
13: gb_est4 : *
14: gb_est5 : *
15: em_estfun : *
16: em_estom : *
17: gb_gss : *
18: em_gss_hum : *
19: em_gss_inv : *
20: em_gss_pln : *
21: em_gss_vrt : *
22: em_gss_fun : *
23: em_gss_mam : *
24: em_gss_mus : *
25: em_gss_other : *
26: em_gss_pro : *
27: em_gss_rtd : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1039	89.4	918	12	BF577781 602092080
2	968.5	83.3	731	13	B1871711 603395825
3	941.5	81.0	728	13	B1870393 603395641
4	882	75.9	777	13	B1819200 603034614
5	861	74.1	828	13	B1596681 603243254
6	861	74.1	834	13	B1766766 603056866
7	857	73.8	561	10	AM763237 ut70d09.y
8	831.5	71.6	948	14	B0707185 AGENCOURT
9	818	70.4	940	14	B0884231 AGENCOURT
10	764	65.7	963	14	B0671259 AGENCOURT
11	754	64.9	1033	11	AK020909 Mus muscu
12	721	62.0	1071	14	BM921213 AGENCOURT
13	703	60.5	951	14	B0674188 AGENCOURT
14	699	60.2	584	10	AM917574 AGENCOURT
15	670	57.7	445	9	AA870722 vq25907.r
16	632.5	54.4	471	9	AA221610 my18d09.r
17	629	54.1	650	12	BG404836 602420160
18	621	53.4	360	10	BE654876 UI-M-BHO-
19	610	52.5	785	13	B1762908 BF041509 BP250025B
20	561.5	48.3	542	12	BF044430 BP250013A
21	549.5	47.3	404	12	BF466521 UI-M-CGDP
22	541	46.6	315	12	BF466521 UI-M-CGDP
23	513	44.1	474	13	B1965174 1d34e07.y
24	506.5	43.6	493	13	BE307031 601087888
25	501.5	43.2	894	13	B1908274 B1908274 603068526
26	486	41.8	558	13	BM484863 538865 MA
27	477	41.0	581	13	B1738634 M03538846
28	473	40.7	538	12	BF821434 MR1-R7003
29	458	39.4	298	14	BM688946 UI-E-C01-
30	429	36.9	278	10	AA417023 52923 MAR
31	428	36.8	342	9	AA637970 vrc3o10.r
32	404.5	34.8	571	12	BF073881 220949 MA
33	404	34.8	345	14	R55379 Y177408.r1
34	392	33.7	587	2	BG686319 602638232
35	373	32.1	400	9	A1152313 ud87h02.r
36	373	32.1	441	13	B1967060 1d27h07.y
37	370	31.8	487	10	AM320117 un17b02.x
38	369.5	31.8	698	13	B1906850 603084633
39	354	30.5	260	10	BB596688 BB596688
40	338	29.1	533	10	BB628951 uu30c03.y
41	336	28.9	910	12	BG110063 602279667
42	335	28.8	456	13	B1966255 1e72904.y
43	316	27.2	687	14	B0208433 UI-R-EP0
44	287	24.7	531	13	B1824443 603038693
45	243.5	21.0	639	10	BB642326 BB642326

ALIGNMENTS

RESULT 1
LOCUS BF577781 918 bp mRNA EST 12-DEC-2000
DEFINITION 602092080F1 NCI_CGAP_C024 Mus musculus cDNA clone IMAGE:4206595 5',
ACCESSION BF577781 mRNA sequence.
VERSION BF577781.1 GI:11651493
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:
Mammalia, Eutheria: Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL COMMENT

Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM9767 row: p column: 20
 High quality sequence start: 17
 High quality sequence stop: 724.
 Location/Qualifiers

FEATURES

1. 918
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone_lib="IMAGE:4206595"
 /lab_host="NCI-CGAP_Co24"
 /note="Organ: colon (71 phage-resistant)"
 Site:2: Salt: Cloned unidirectionally. Primer: Oligo dT.
 Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCI-CGAP Library."
 BASE COUNT 153 a 292 c 268 g 204 t 1 others
 ORIGIN

Alignment Scores:

Pred. No.: 2,04e-99 Length: 918
 Score: 1039.00 Matches: 214
 Percent Similarity: 98.17% Conservative: 0
 Best Local Similarity: 98.17% Mismatches: 4
 Query Match: 89.41% Indels: 3
 DB: 12 Gaps: 0

US-09-245-198a-2 (1-225) x BF577781 (1-918)

OY 8 LeuAlaCysLeuGlyLeuLeuLeuValValSerLeuGlySerTPaJarThrLeuSer 27
 |||||
 DB 2 CTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 61
 OY 28 AlAGInGluProSerGInGInGluLeuThrAlAGInAspArgArgGluProProGInLeu 47
 |||||
 DB 62 GCCAGAGAGGCTCTCGAGGAGAGCTGACAGACAGACGCGCGGAGCCCGCTGAAGTG 121
 OY 48 AspProGInThrGInGInSerGInAspValValProPheLeuGInGInLeuValArgPro 67
 |||||
 DB 122 AATCCCAAGACAGAGAAAGCAGAGATGTGTACTTTTGGACAACTAATGTCGGCT 181
 OY 68 ArgArgSerAlaProGlyGlyArgGlyAlaArgProArgArgAlaAlaAlaHisTyr 87
 |||||
 DB 182 CGAAGAGAGTGTCTTAAGCGCGGAGCGGCGCT -CGCGAGAGTATTTGACAGCCATTA 240
 OY 88 GluValHisProArgProGInGInAspGlyAlAGInAlaGlyValAspGlyThrValSer 107
 |||||
 DB 241 GAGGTTATCTCGCGGAGAGAGATGAGACAAACAGATGTGTGAGTGGACAGTAGT 300
 OY 108 G1YTPGInGluThrGlyHisLeuSerSerProLeuArgTyrAspArgGInIlleGly 127
 |||||
 DB 301 GCGTGGGAAGAGCAAAATCAACAGCTCCAGCCCTGTGGCTACGACCCCGCAGATTGGG 360
 OY 128 GluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGInValHisPheAspGlu 147
 |||||
 DB 361 GAATTTCAGTCAATCAGGCTGGGCTCTACTACTGTAAGTGTGAGTGCATTGTGATGAG 420
 OY 148 G1YValAlaValTyrLeuGlyLeuAspLeuLeuValAsnGlyValIleuAlaLeuArgCys 167
 |||||
 DB 421 GGAAGAGCTGTACCTGAGAGCTGTGCTGTGTAACGGTGTGCGCTGTGCTGCTGC 480
 OY 168 LeuGInGluPheSerAlaThrAlaAlaSerSerProGlyProGInLeuAspGluCysGIn 187
 |||||
 DB 481 CTGGAAAGATTTCTAGCCACAGACAGAAAGCTTCCTGGGCGCCAGCTCCGTTGTGTCCAG 540

OY 188 ValSerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeuProTyr 207
 |||||
 DB 541 GTGTCTGGGCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 599
 OY 208 AlaHisLeuValAlaAlaProPheLeuThrTyrPheGlyLeuPheGInValHis 225
 |||||
 DB 600 GCTTCATCTTAAGGCTGC -CCCTTCCTAAGTAACTTGGAGCTCTTCAAGTTCAC 652

RESULT 2

BI871711

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Human sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 1 (bases 1 to 731)
 NIH-MGC <http://imgc.ncl.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Plate: LLAM12034 row: b column: 07
 High quality sequence stop: 728.
 Location/Qualifiers

FEATURES

source

1. 731
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:5405478"
 /clone_lib="NIH-MGC-90"
 /tissue_type="adenoecarcinoma, cell line"
 /lab_host="DH10b (phage-resistant)"
 /note="Organ: liver; Vector: PCMV-SPORT6; Site:1: NotI; Site:2: Salt: Cloned unidirectionally; oligo-dT primed. Average insert size 1.7 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH-MGC Library."
 BASE COUNT 125 a 242 c 228 g 136 t
 ORIGIN

Alignment Scores:

Pred. No.: 4.19e-92 Length: 731
 Score: 968.50 Matches: 195
 Percent Similarity: 92.31% Conservative: 9
 Best Local Similarity: 88.24% Mismatches: 16
 Query Match: 83.35% Indels: 2
 DB: 13 Gaps: 1

US-09-245-198a-2 (1-225) x BI871711 (1-731)

OY 6 LeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValSerLeuGlySerTPaJarThr 25
 |||||
 DB 2 CTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 61
 OY 26 LeuSerAla---GInGluProSerGInGInGluLeuThrAlAGInAspArgArgGluPro 44
 |||||
 DB 62 CTGTCCGCCAGCAGAGAGCTGTGCCAGAGAGCTGTGTGCAAGAGAGACCGAC 121
 OY 45 ProGInLeuAsnProGInThrGInGInSerGInAspValValProPheLeuGInGInLeu 64
 |||||
 DB 122 TCAGAACTGAATCCCGCAGACAGAAAGCCAGAGATCTCGGCTTTCGTAACCGAGTA 181

QY 46 GlutenaSPROGlnThrGluSerGlnAspValValProphLeuGluGlnLeuVal 65
|||||
Db 44 GAACGAATCCCGACAGCAAGAAAGCCAGATCCCTGCGCTTCGTAACCGACTACTT 103
QY 66 ArgProArGArSerAlaProLysGlyArgLysAlaValProArGArGAlaIleAlaAla 85
104 CGGCTCGCAAGATGCCCTTAAGGCCGGAACACCGGGCTCGAAGCGCATCCACCC 163
QY 86 HisTyrGluValHisProArGProGlyGlnAspGlyAlaGlnAlaGlyValAspGlyThr 105
164 CATATGAAATTCATCCACGACCTGACAGACGAGGCGCAGCAGGAGTGTGACGAGCACA 223
QY 106 ValSerGlyTTPGluGluThrLysIleAsnSerSerProLeuArGTYrAspArgGln 125
224 TTGAGTGGCTGGAGGAGCAAGATCAACAGCTCCACCTCTGCTACACCGCCAG 283
QY 126 IleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPhe 145
284 ATCGGGGAGTTTATAGTCACCCGGGCTGGCTACTACTACTGTCAGGTGACTTT 343
QY 146 AspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGlyValLeuAlaLeu 165
344 GATGAGGGGAGGCTGTCTACCTGAGCTGACTGCTGAGTGTGCTGCTGCTG 403
QY 166 ArgCysLeuGluGluPheSerAlaThrAlaIleSerSerProGlyProGlnLeuArgLeu 185
404 CGCTGGCTGGAGAAATTCACGACCTGCGCCAGCTCCGCGGCCACCTCCGCTC 463
QY 186 CysGlnValSerGlyLeuLeuProLeuArGProGlySerSerLeuArgIleArGThrLeu 205
464 TGCCAGAGTGTCTGGCTGTGGCCCTGCGGCGCAGGCTCTCCGCGCATCCGACCTC 523
QY 206 ProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225
524 CCTGGGCGCCATCTCAAGGCTGCCCTCTCTCCTACCTTGAGCTCTCCAGGTTTAC 583
RESULT 6 834 bp mRNA Linear EST 25-SEP-2001
LOCUS B1766766
DEFINITION 60305686F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5206217 5',
mRNA sequence.
ACCESSION B1766766
VERSION B1766766.1 GI:15758344
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 834)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1A11517 row: c column: 18
High quality sequence stop: 772.
Location/Qualifiers
1. 834
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5206217"
/clone_1db="NIH_MGC_122"
/lab_host="DH10B"
/note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
Site_1: NotI; Site_2: EcoRV (destroyed); RNA source

anonymous pool of 24 week female lung, 16 week female
spleen, and 20-22 week male spleens. Library is oligo-dT
primed and directionally cloned (EcoRV site is destroyed
upon cloning). Average insert size 1.4 kb, insert size
range 1-3 kb. Library is normalized and enriched for
full-length clones and was constructed by C. Gruber
(Invitrogen). Research Genetics tracking code 026. Note:
this is a NIH_MGC library."

BASE COUNT 154 a 287 c 223 g 170 t
ORIGIN

Alignment Scores:
Pred. No.: 1 19e-80 Length: 834
Score: 861.00 Matches: 172
Percent Similarity: 92.27% Conservative: 7
Best Local Similarity: 88.66% Mismatches: 14
Query Match: 74.10% Indels: 2
DB: 13 Gaps: 0

US-09-245-198a-2 (1-225) x B1766766 (1-834)

QY 33 GlnGluGluLeuThrAlaGluAspArgArgGluProGluLeuAsnProGlnThrGlu 52
|||||
Db 3 CAGAGGAGCTGTGGGACAGAGAGACAGACCGCTGCGAATCCACAGACAGAA 62
QY 53 GluSerGlnAspValAlaProPheLeuGluGlnLeuValArgProArGArSerAlaPro 72
|||||
Db 63 GAAAGCCAGATCTCCGCTTCTTCTGAACCGCACTGACTGCGCTCGCAGAAAGTGCACCT 122
QY 73 LysGlyArgLysAlaArgProArGArGAlaIleAlaHisTyrGluValHisProArG 92
123 AAAGCGCGGAAACACAGCGGCTCGAAGAGCATCCAGCCCATTTGAAAGTTCATCCACGA 182
QY 93 Pro-GlyGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPglGluThr 112
183 CCGTGGACAGAGCAGGAGCGCAGGAGGTGTGACGGGACAGTGAAGTGTGGAGGAAG 242
QY 112 LysIleAsnSerSerSerProLeuArGTYrAspArgGlnIleGlyGluPheThrVal 132
243 CAGATCAACAGCTTCACAGCCCTTGCCTGCTACACACCGCCAGTATAGTTCAC 302
QY 132 eArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyr 152
303 CCGGAGTGGCTCTACTACTGCTGACTGTCAGAGTGTGATGAGGGAAGGCTGTCTA 362
QY 152 rLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLeuGluGluPhe 172
363 CCTGAAGCTGAGCTTGTGGTGTGATGTGTGCTGGCCCTGCGCTGGAGGAATCTTC 422
QY 172 rAlaThrAlaIleSerSerProGlyProGlnLeuArgLeuGlyGlnValSerGlyLeu 192
423 AGCCACTGCGGCGCAGTTCCTCGGGGCCCGCCAGCTCGCTGCGCAGAGTGTGGGCTT 482
QY 192 uProLeuArGProGlySerSerLeuArGlyIleArgThrLeuProTrpAlaHisLeuLysAl 212
483 GGCCCTGGGCGCAGATGTC-TCCTGGGAGATCCGACACCTCCCTGGGCGCATCAAGGC 541
QY 212 aAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 225
542 TGCCTCCCTTCTCTCACTTGGAGCTTCCAGGTTTAC 581

RESULT 7 561 bp mRNA Linear EST 04-MAY-2000
LOCUS AM763237
DEFINITION ut70009.y1 NCI_CGAP_Mem3 Mus musculus cDNA clone IMAGE:315633 5',
similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS
/, mRNA sequence.
ACCESSION AM763237
VERSION AM763237.1 GI:7695174
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 561)
 AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 JOURNAL Tumor Gene Index
 COMMENT Unpublished (1997)
 Other ESTs: ur70d09.x1
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Lohar Hennighausen Ph.D., Chu Xia Deng Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/resources.shtml

MG1:1058389
 Seq primer: -40RP from G1bco
 High quality sequence stop: 433.
 Location/Qualifiers
 1..561
 /organism="Mus musculus"
 /strain="129,C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:3155633"
 /clone_lib="NCI-CGAP Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;
 Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lohar Hennighausen/Chu Xia Deng, NIH
 Reference for transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."

BASE COUNT 108 a 158 c 194 g 100 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1 77e-80 Length: 561
 Score: 857.00 Matches: 167
 Percent Similarity: 99.40% Conservative: 0
 Best Local Similarity: 99.40% Mismatches: 1
 Query Match: 73.75% Indels: 0
 DB: 10 Gaps: 0

US-09-245-198a-2 (1-225) x AM763237 (1-561)

OY 1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeu 20
 |||||
 DB 58 GTGCTGAGCTGGGCGGCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 117
 OY 21 GlySerThrAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluASP 40
 |||||
 DB 118 GGGAGCTGGGCAACGCTGTCGCCAGAGACCTTCAGAGAGAGCTGACAGAGAGAC 177
 OY 41 ArgArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPhe 60
 |||||
 DB 178 CGCGGGAGAGCCCTGACATGATCCCGACAGAGAAAGCCAGATGCGTACCTTTC 237
 OY 61 LeuGluGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArg 80
 |||||
 DB 238 TTGGAACAACACTAGTCGCGCTCGAAGAGTCTCTTAAGAGCGCGAAGCGCGGCTCGC 297
 OY 81 ArgAlaIleAlaIleAlaHsTrpGluValHisProArgProGlyGlnAspGlyAlaGlnAla 100
 |||||
 DB 298 CGAGCAATTCCACCCCACTTGTGAGTTCATCTCGGCCAGACAGATGAGCAACAAGCA 357
 OY 101 GlyValAspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerProLeu 120
 |||||
 DB 358 GGTGTGATGGGAGAGTGAAGTGGTGGAGAGACCAAAATCAACAGCTCCACCCCTCTG 417

OY 121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTrpLeuTyr 140
 |||||
 DB 418 CGCTACGACCGCCAGATGTGGGGAATGTACATCATCAGGCGGCGCTACTACCTCTAC 477
 OY 141 CysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsn 160
 |||||
 DB 478 TGTACAGTGCACCTTGTATGATGGGAAAGCGTCTTACTCTGAAGCTGATGTAAC 537

OY 161 GlyValLeuAlaLeuArgCysLeu 168
 |||||
 DB 538 GGTGTGCTGGCCCTGGCTGCTCG 561

RESULT 8
 LOCUS BQ707185
 DEFINITION AGENCOURT_8353983 NIH_MGC_113 Homo sapiens CDNA clone IMAGE:6278608
 5', mRNA sequence.
 ACCESSION BQ707185
 VERSION BQ707185
 KEYWORDS BQ707185.1 GI:21846084
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 948)
 AUTHORS NIH-MGC <http://imgc.ncbi.nlm.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. Mark Watson
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LCM2466 row: n column: 17
 High quality sequence start: 24
 High quality sequence stop: 550.
 Location/Qualifiers

1..948
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6278608"
 /clone_lib="NIH_MGC_113"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; CDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GCCACGAG(G). Library constructed by Ling Hong in the
 laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC Library."

FEATURES
 source

BASE COUNT 173 a 334 c 252 g 187 t 2 others
 ORIGIN

Alignment Scores:
 Pred. No.: 1 9e-77 Length: 948
 Score: 831.50 Matches: 164
 Percent Similarity: 92.31% Conservative: 4
 Best Local Similarity: 90.11% Mismatches: 13
 Query Match: 71.56% Indels: 2
 DB: 14 Gaps: 1

US-09-245-198a-2 (1-225) x BQ707185 (1-948)

OY 44 ProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeuGluGln 63
 |||||
 DB 3 CCCTCGGAAGTGAATCCCGACAGAGAAAGCCAGAGATCTCGGCTTCCGTAACCA 62
 OY 64 LeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgAlaIle 83

PUBMED
REFERENCE
AUTHORS

11217851
5 (bases 1 to 1033)
Aachchi,J., Aizawa,K., Akahira,S., Akimura,T., Aono,H., Arai,A.,
Arikawa,T., Baldarelli,R., Bono,H., Brownstein,M., Bull,C.,
Carninci,P., Fukuda,S., Fukunishi,Y., Furuno,M., Hanagaki,T.,
Hara,A., Hayatsu,N., Hill,D., Hiramoto,K., Hirooka,T., Horii,F.,
Hume,D., Imotani,K., Ishii,Y., Itoh,M., Izawa,M., Kasukawa,T.,
Kato,H., Kawai,J., Kojima,Y., Konno,H., Kouda,M., Koya,S.,
Kurihara,C., Matsuyama,T., Miyazaki,A., Nishi,K., Nomura,K.,
Numazaki,R., Ohno,M., Okazaki,Y., Okido,T., Owa,C., Quackenbush,J.,
Saito,H., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D.,
Schirml,L., Shibata,K., Shibata,Y., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F.,
Tanaka,T., Tejima,Y., Toya,T., Yamamura,T., Yamataka,I.,
Yanunishi,A., Yoshida,K., Yoshino,M., Muramatsu,M., and
Hayashizaki,Y.

Direct Submission
Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suheiho-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)

COMMENT
Please visit our web site (<http://genome.gsc.riken.go.jp/>) for
further details.

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues. First strand cDNA was primed with a primer
[5' GAGACAGACAGACGATCCAGACGCTCTTTTCTTTTCTTTTCTTNN 3'], cDNA was
prepared by using trehalose thermo-activated reverse transcriptase
and subsequently enriched for full-length by cap-trapper. cDNA went
through two rounds of normalization to Rot = 20.0 and subtraction
to Rot = 458.8. Second strand cDNA was prepared with the primer
adaptor of sequence [5'
GAGACAGAGATTTTCGAGTAAATTAATTAATTCCTCCCCCCCC 3']. cDNA was cleaved
with BamHI and XhoI. Vector: a modified pluescript KS(+) after
bulk excision from Lambda FLC I. Cloning sites, 5' end: SalI; 3'
end: BamHI. Host: DH10B. Retina RNA was provided by Stefano
Gustcinich (Department of Neurobiology, Harvard Medical School, 220
Longwood Ave., Boston, MA 02115, USA) whose assistance is
gratefully acknowledged.

Retina RNA was provided by Stefano Gustcinich (Department of
Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston,
MA 02115, USA) whose assistance is gratefully acknowledged.

FEATURES
SOURCE
location/Qualifiers
1..1033
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM.DB:A930030D13"
/db_xref="MGD:MGI:1191867"
/db_xref="taxon:10090"
/clone="A930030D13"
/tissue_type="retina"
/clone_id="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
1..1033
/gene="Tnfrsf12"
1..433
/gene="Tnfrsf12"
/note="data source:MGD, source key:MGI:1196259,
evidence:ISS
putative
tumor necrosis factor (ligand) superfamily, member 12"
/codon_start=2
/protein_id="BAB32249.1"
/db_xref="GI:12861641"
/db_xref="MGD:MGI:1196259"
/translation="IAAHYEYHPRPGDGAQAGVDGVSCHEEKINSSPLRDROI
GEFVYIAGLTLYCYGVHFBDGKAVYLAKLIDLVNGVLAFLVGLTGLFQVH"
LCQVSGILPRPGSSLRIRLPLPMHLLAFLDVLVGLTGLFQVH"

BASE COUNT	228 a	300 c	248 g	256 t	1 others
ORIGIN					
Alignment Scores:					
Pred. No.:	3,4e-69				1033
Score:	754.00				Matches: 143
Percent Similarity:	100.00%				Conservative: 0
Best Local Similarity:	100.00%				Mismatches: 0
Query Match:	64.89%				Indels: 0
DB:	11				Gaps: 0
US-09-245-198A-2 (1-225) x AK020909 (1-1033)					
OY	83	11e1a1a1a1h1styg1g1u1a1h1s1p1o1a1r1p1o1g1i1n1a1s1p1g1y1a1g1n1a1g1y1a1	102		
Db	2	ATTGACGCCCATTTAGAGAGTTTCATCTCGGCCAGGACAGATGAGACAAAGCAGTGTG	61		
OY	103	ASpGLyThrValSerGLyTTrpGLuGLuThrLysIleAsnSerSerProLeuArgTyr	122		
Db	62	GATGGAGACAGTGAAGTGGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTTCGGCTAC	121		
OY	123	ASpArgGlnIleGLyGLuPheThrValIleArgAlaGLyLeuTyrTyrLeuTyrCysGln	142		
Db	122	GACCCGAGATTGGGGAATTTACAGTCATCAAGGCTGGGCTCTACTACCTGACTGTGAC	181		
OY	143	ValHisPheAspGLuGLyLysAlaValIlyrIleuLysLysAspLeuValAsnGLyVal	162		
Db	182	GTGCACCTTGTGATGAGAGGAAGGCTCTACCTGAGAGCTGACCTGCTGTAACGTTGTG	241		
OY	183	LeuAlaIleuArgCysLeuGLuGLuPheSerAlaThrAlaIleSerSerProGLyProGln	182		
Db	242	CTGGCCCTCGCGCTGCTGTGAGAAATTTCTACGACACAGACAGCAAGCTCTCTGGCCCCAG	301		
OY	183	LeuArgLeuCysGlnValSerGLyLeuLeuProLeuArgProGLySerSerLeuArgIle	202		
Db	302	CTCCGTTTGTGGCAGGTCTGTGGCTGTGGCGCTGGCGGCGAGGGTCTTCCCTTGGATTC	361		
OY	203	ArgThrLeuProGTPAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGLyLeuPhe	222		
Db	362	CGCACCCCTCCCTGGGCGTCATCTTAAGGCTGCCCCCTTCTTAACCTACTTTGGACTTTT	421		
OY	223	GlnValHis 225			
Db	422	CAAGTTCAC 430			
RESULT 12					
BM921213					
LOCUS	BM921213	1071 bp	mRNA	linear	EST 12-MAR-2002
DEFINITION	AGNCNCOURT_6633046 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752561				
ACCESSION	5', mRNA sequence.				
VERSION	BM921213				
KEYWORDS	BM921213.1 GI:19371592				
SOURCE	EST.				
ORGANISM	human.				
REFERENCE	Homo sapiens				
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
JOURNAL	1 (bases 1 to 1071)				
COMMENT	NIH-MGC http://mgc.nci.nih.gov/				
	unpublished (1999)				
	National Institutes of Health, Mammalian Gene Collection (MGC)				
	Contact: Robert Strausberg, Ph.D.				
	Email: cgaabs-r@mail.nih.gov				
	Tissue Procurement: Life Technologies, Inc.				
	cDNA Library Preparation: Life Technologies, Inc.				
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)				
	DNA Sequencing by: Agencourt Bioscience Corporation				
	Clone distribution: MGC clone distribution information can be				
	found through the I.M.A.G.E. Consortium/LNL at:				
	http://image.lnl.gov				
	Plate: LLM12786	row: p	column: 02		
	High quality sequence stop: 656.				
	Location/Qualifiers				
FEATURES					

RESULT 14	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES
422	CTCCCGGGGCCCATTTAAAGGCGCCCGCTTTCCTACCATTTTGGAAATTTT 477	584 bp mRNA linear EST 25-MAY-2000	AW917574	AW917574	GI:8083328	EST.	Rattus norvegicus	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.	1 (bases 1 to 584)	Lee,N.H., Glodex,A., Chandra,I., Mason,T.M., Quackenbush,J., Kerlavage,A.R. and Adams,M.D.	Rat Genome Project: Generation of a Rat EST (REBT) Catalog & Rat Gene Index	Unpublished (1998)	Location/Qualifiers
													1. 584
													/organism="Rattus norvegicus"
													/db_xref="taxon:10116"
													/clone="RCIEF49"
													/clone_lib="Rat gene index, normalized rat, norvegicus, Bento Soares"
													/tissue_type="mix - brain, ovary, placenta, kidney, lung, liver, embryo, heart, muscle, spleen"
													/lab_host="SOLR"
													/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Estimated insert size approx.1 kb"
													BASE COUNT 108 a 188 c 147 g 138 t 3 others
													ORIGIN
													Alignment Scores:
													Pred. No.: 9.72e-64
													Score: 699.00
													Matches: 133
													Best Local Similarity: 99.26%
													Conservative: 1
													Mismatches: 1
													Query Match: 60.15%
													Indels: 0
													DB: 10 Gaps: 0
													US-09-245-198A-2 (1-225) x AW917574 (1-584)
													QY 91 ProArgProGlyGlnAspGlyAlaGlnAlaGlyAlaAspGlyThrValSerGlyTyrGln 110
													Db 2 CCACAGCCAGCAGACAGATGAGACACAGCAGAGTGTGATGGACAGTGAAGTGGGAA 61
													QY 111 GluThrLysIleAsnSerSerSerProLeuArgTyrAspArgGlnIleGlyGluPheThr 130
													Db 62 GAGACCAAAATCAACAGCTCCAGCCCTCTGGCTATGACCCGCAATGGGGGAATTTACG 121
													QY 131 ValIleArgAlaGlyLysLeuTyrTyrLeuTyrCysGlnValHisPheAspGlnGlyLysAla 150
													Db 122 GTCATCAGGCGCTGGCTTACTACTGTACTGTACAGGTGACATTTCAGTGAAGGGGAAAGCA 181
													QY 151 ValTyrLeuLysLeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLeuGlnGlu 170
													Db 182 GTCTACCGTAGAGCTGCAGCTTGTGTGAATGATGTGCTGGCCCTGCGTGGTGAAGA 241
													QY 171 PheSerAlaThrAlaIleAsnSerSerProGlyProGlnLeuArgLeuCysGlnValSerGly 190

Db	242	TTCTACGCACAGCAGCGACTCTCTGGCCCAAGCTCCGTCGTGGCAGGTGCTGGG	301
Oy	191	IeuLeuProLeuAryProGlYSerSerLeuAryIleArGrHrLeuProTPAlaHisLeu	210
Db	302	CTGTGCCCTMGCGGCGCAGGGGTCTTCCCTCGGATCGGACCCCTCCCTGGGCTCATCTT	361
Oy	211	LysAlaIalProPheLeuThrTyrrPheGlyLeuPheGlnValHis	225
Db	362	AAGGCTGCCCTTCCTTACTTGTGACTCTMTCAAGTTTCAC	406
RESULT 15			
LOCUS	AA870722	445 bp	mRNA linear EST 16-MAR-1998
DEFINITION	vq25g97.f1 Barsstead stromal cell line MFLRB8 Mus musculus cDNA		
ACCESSION	AA870722		
VERSION	AA870722.1	G1:2966167	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 445)		
AUTHORS	Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T., Geisell,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M., Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B., Thaising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and Waterson,R.		
TITLE	The WashU-HIMI Mouse EST Project		
JOURNAL	Unpublished (1996)		
COMMENT	Contact: Marra M/Mouse EST Project WashU-HIMI Mouse EST Project Washington University School of Medicine# 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108 Tel: 314 286 1800 Fax: 314 286 1810 Email: mouseest@wustl.edu This clone is available royalty-free through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov) for further information. MGI:601556 Seq primer: -28m13 rev2 ET from Amersham High quality sequence stop: 420.		
FEATURES			
source	Location/Qualifiers		
	1..445		
	/organism="Mus musculus"		
	/db_xref="taxon:10090"		
	/clone_image="IMAGE:1095324"		
	/clone_lib="Barsstead stromal cell line MFLRB8"		
	/cell_line="C2C12 (undifferentiated)"		
	/lab_host="DH10B"		
	/note="Vector: pRT73D-Pac (Pharmacia) with a modified polylinker. Site_1: EcoRI; Site_2: NotI; 1st strand cDNA was primed with a Not I - oligo(dt) primer [5' TGTAACAATCTGAAGTGGAGGCGGCGCCCTTTTTTTTTTTTTTTTTT 3']; double-stranded cDNA was ligated to Eco RI adaptors [AATTCGATCCCTGG], digested with Not I and cloned into the Not I and Eco RI sites of the modified pRT73 vector. Source undifferentiated tissue culture cell line C2C12. Library constructed by Bob Barsstead. The C2C12 cell line (available from ATCC, catalog # CRL-1772) differentiates rapidly, forming contractile myotubes and producing characteristic muscle proteins."		
BASE COUNT	102 a 119 c 135 g 89 t		
ORIGIN			
Alignment Scores:			
Pred. NO.:	7.64e-61	Length:	445
Score:	670.00	Matches:	132
Percent Similarity:	96.38%	Conservative:	1
Best Local Similarity:	95.65%	Mismatches:	5
Query Match:	57.66%	Indels:	1
DB:	9	Gaps:	0

GenCore version 5.1.4-p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OK protein - protein search, using sw model

Run on: March 31, 2003, 07:42:57 ; Search time 11.9352 Seconds

(without alignments)
781.906 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162

Sequence: 1 VLSTGLALACILGLLVVSL.....PMAHLKAPFLTYGFLFOVH 225

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1162	100.0	225	1	054907 mus musculus
2	1020	87.8	249	1	043508 homo sapien
3	109.5	9.4	272	1	091868 gallus galli
4	102.5	8.8	316	1	035235 m tumor nec
5	90.5	7.8	260	1	097626 canis famli
6	90.5	7.8	318	1	095822 r tumor nec
7	88	7.6	332	1	096643 homo sapien
8	87	7.5	244	1	096dm3 actus trivi
9	87	7.5	261	1	096dm3 callithrix
10	87	7.5	261	1	043557 homo sapien
11	86.5	7.4	240	1	043557 homo sapien
12	86.5	7.4	1237	1	043557 homo sapien
13	86	7.4	261	1	043557 homo sapien
14	86	7.4	261	1	043557 homo sapien
15	85	7.3	240	1	043557 homo sapien
16	85	7.3	1237	1	043557 homo sapien
17	84	7.2	278	1	043557 homo sapien
18	83.5	7.2	241	1	043557 homo sapien
19	83.5	7.2	250	1	043557 homo sapien
20	82	7.1	530	1	043557 homo sapien
21	82	7.1	535	1	043557 homo sapien
22	82	7.1	535	1	043557 homo sapien
23	80.5	6.9	246	1	043557 homo sapien
24	80.5	6.9	1234	1	043557 homo sapien
25	79.5	6.8	197	1	043557 homo sapien
26	79	6.8	920	1	043557 homo sapien
27	79	6.8	1584	1	043557 homo sapien
28	79	6.8	2468	1	043557 homo sapien
29	78.5	6.8	260	1	043557 homo sapien
30	78.5	6.8	1164	1	043557 homo sapien
31	78	6.7	285	1	043557 homo sapien
32	78	6.7	763	1	043557 homo sapien
33	77.5	6.7	201	1	043557 homo sapien

34	77.5	6.7	261	1	TNFS_PIG	095mq5 sus scrofa
35	77.5	6.7	787	1	NASP_HUMAN	P49321 homo sapien
36	77	6.6	279	1	TNFS_MOUSE	P41047 mus musculus
37	77	6.6	817	1	NAHL_BOVIN	Q28036 bos taurus
38	76.5	6.6	214	1	SMP_ECOLI	P18838 escherichia
39	76	6.5	695	1	APP2_MOUSE	006333 mus musculus
40	75.5	6.5	250	1	TN13_HUMAN	075888 homo sapien
41	75.5	6.5	575	1	MIS_BOVIN	P03972 bos taurus
42	75.5	6.5	999	1	ORXP_RAT	063617 rattus norv
43	75.5	6.5	1058	1	UBA1_HUMAN	P22314 homo sapien
44	75	6.5	1004	1	SAL2_MOUSE	Q9q96 mus musculus
45	75	6.5	3305	1	APLP_MANSE	Q25490 manduca sex

ALIGNMENTS

RESULT 1	ID	TN12_MOUSE	STANDARD:	PRT:	225 AA.
AC	054907	09CTP2:			
DT	15-JUN-2002	(Rel. 41, Created)			
DT	15-JUN-2002	(Rel. 41, Last sequence update)			
DE	15-JUN-2002	(Rel. 41, Last annotation update)			
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (Fragment).				
CN	TNFSF12.				
OS	Mus musculus (Mouse).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX	NCBI_TaxID=10090;				
OX	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Peritoneal macrophage;				
RC	MEDLINE=98070415; PubMed=9405449;				
RA	Chicopeurche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,				
RA	Hession C., Garcia I., Browning J.L.;				
RT	"TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis."				
RT	J. Biol. Chem. 272:32401-32410(1997).				
RN	[2]				
RP	SEQUENCE OF 83-225 FROM N.A.				
RC	STRAIN=C57BL/6J; TISSUE=Retina;				
RC	MEDLINE=21085660; PubMed=11217851;				
RA	Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,				
RA	Arakawa T., Hara A., Fukunishi Y., Kono H., Aichi J., Fukuda S.,				
RA	Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,				
RA	Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,				
RA	Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,				
RA	Fleischmann W., Gaasterland T., Glass C., King B., Kochiwa H.,				
RA	Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,				
RA	Schirini L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,				
RA	Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,				
RA	Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,				
RA	Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,				
RA	Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,				
RA	Lyons P., Marchionni L., Mashima J., Mazzarelli J., Momberts P.,				
RA	Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,				
RA	Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,				
RA	Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Winding L.,				
RA	Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohitsuki S.,				
RA	Hayashizaki Y.;				
RL	"Functional annotation of a full-length mouse cDNA collection."				
RL	Nature 409:685-690(2001).				
CC	-I- FUNCTION: Binds to FN14 and possibly also to TNFSF12/ADP3. Weak inducer of apoptosis in some cell types. Promotes angiogenesis and the proliferation of endothelial cells. Mediates NF-kappaB activation (By similarity).				
CC	-I- SUBUNIT: Homotrimer (potential).				
CC	-I- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By similarity).				
CC	-I- TISSUE SPECIFICITY: WIDELY EXPRESSED.				
CC	-I- PTM: The soluble form is produced from the membrane form by				

	Query Match	87.88:	Score 1020:	DB 1:	Length 249:
	Best Local Similarity	88.88:	Pred. No. 2.6e-84:		
	Matches 199:	Conservative	9:	Mismatches 16:	Indels 0:
				Gaps	0:
QY	2	LSTGLIATLACGLDLLVVSLSGSMATLSAQPSEOEELTAARDREPPRLANQTESQDPVEFL	61		
Db	26	LGLIATLACGLDLLLVVSLSGSRASLSAQEPQEEELVAEEDQPSLNLQTESQDPAPFL	85		
QY	62	EOLYRPRRSAPKGRKARRRRLAAHYEVNHPRGODGAAGVDGYTSGMEETKINSSPLR	121		
Db	86	NRLYRPRRSAPKGRKTRARRRLAAHYEVNHPRGODGAAGVDGYTSGMEETKINSSPLR	145		
QY	122	YDRQIGFEFTYIRAGLYLYLYCVNHFEDEGKAVYLKLDLNVGLALRCLEEFSTAASDPG	181		
Db	146	YNROIQEFTYIRAGLYLYLYCVNHFEDEGKAVYLKLDLVDVGLALRCLEEFSTAASSLGP	205		
QY	182	QLRLQVSGILRLPRGSSLRIRTLPLMAILKAPAPFLTYGGLQVNH	225		
Db	206	QLRLQVSGILRLPRGSSLRIRTLPLMAILKAPAPFLTYGGLQVNH	249		

RESULT 3			
ID	TNFS_CHICK	STANDARD:	PRT: 272 AA.
AC	Q918D8;		
DT	15-JUN-2002 (Rel. 41, Created)		
DT	15-JUN-2002 (Rel. 41, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-L) (CD154 protein).		
GN	TNFS5 OR CD40LG OR CD40L.		
OS	Gallus gallus (Chicken).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus.		
OX	NCBI_TaxID=9031;		
RN	[1]		
RF	SEQUENCE FROM N.A.		
RC	STRAIN-White leghorn; TISSUE-Spleen;		
RA	Tregaskes C.A., Young J.R., Burnside J.;		
RL	Submitted (JUN-1999) to the EMBL/Genbank/DBJ databases.		
CC	-1- FUNCTION: Cytokine that binds to TNFRSF5. Mediates B-cell proliferation in the absence of co-stimulus as well as IgE production in the presence of IL-4. Involved in Immunoglobulin class switching (By similarity).		
CC	-1- SUBUNIT: Homotrimer (By similarity).		
CC	-1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an extracellular soluble form (By similarity).		
CC	-1- PFM: The soluble form derives from the membrane form by proteolytic processing (By similarity).		
CC	-1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.		
CC	-----		
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/annouce/ or send an email to license@isb-sib.ch).		
CC	-----		
DR	EMBL: AJ243435; CAB95748.1; -		
DR	HSSP: P29965; IALY		
DR	InterPro: IPR003263; TNF_5.		
DR	InterPro: IPR003636; TNF_abc.		
DR	InterPro: IPR000478; TNF_family.		
DR	Pfam: PF00229; TNF_1.		
DR	ProDom: PD002012; TNF_abc; 1.		
DR	ProDom: PD008600; TNF_5; 1.		
DR	SMART: SM00207; TNF_1.		
DR	PROSITE: PS00251; TNF_1; 1.		
DR	PROSITE: PS00049; TNF_2; 1.		
DR	Cytokine; Transmembrane; Glycoprotein; Signal-anchor.		
TM	CHAIN 1 272		
	TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY		

			MEMBER 5, MEMBRANE FORM, TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY MEMBER 5, SOLUBLE FORM (BY SIMILARITY). CYTOPLASMIC (POTENTIAL). SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN) (EXTRACELLULAR).	
FT	CHAIN	111	272	
FT	DOMAIN	1	23	
FT	TRANSMEM	24	44	
FT	DOMAIN	45	272	
FT	SITE	110	111	CLEAVAGE (BY SIMILARITY). POTENTIAL.
FT	DISULFYD	190	229	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBONHYD	124	124	N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBONHD	146	146	N-LINKED (GLCNAC. .) (POTENTIAL).
SO	CARBONDH	251	251	
	SEQUENCE	272 AA;	30862 MW; 3409254ABE53CDD CRC64;	
Query Match		9.4%;	Score 109.5; DB 1; Length 272;	
Best Local Similarity		22.5%;	Pred. No. 0.012;	
Matches		Conservative 58;	Mismatches 107; Indels 53; Gaps 10;	
OY	I VLSUGLALACLGLLVVSLGSMATLSAO-----EPSQBELTAEDRRP-----P	45		
Db	: : : I : : : : : :	:		
OY	34 VOITISTVFCLTYLIHKMKMKEEVLSIANDDYIFLRKVKQCQGDEDKSTLLDCERVLKGFO	93		
OY	46 ELNPOTESOQDVVPRLVDLPKRKSAPRGGR-----ARRRAIAAHYEHNRPQDGQAQ	100		
Db	: : : : : : : : : : : : :	:		
OY	94 DLQQCDRTASEELPKFFENHRSHENDPHLSRNETSVAEEKRPIAIHTLA-----GVKS	145		
OY	101 GVDTGSVCMEETK-INSSSPRLRYDRQIQEFNYIRAGLYLVLCOVHFDESKA-----V	151		
Db	: : : : : : : : : : : :	:		
OY	146 NTTVAYALKWMTTSAPTSSLISIYNH--GKLVEKKAGLYTIYSQSFCIKAAASAFPTLYI	203		
OY	152 YLKIDLNLNGVALRCIEESFATSAASSPPQLRIQCV-----SGILLPRSSSLRIPTLP	206		
Db	: : : : : : : : : : : :	:		
OY	204 XLYIPMEDRDLL-MKGLDIHSYSTA-----LCELDINSREGVGVELRKQGDIVNVNTD	254		
OY	207 WAHLKAAPFLYYFGLFQV	224		
Db	: : : : : : : : : : : :	:		
OY	255 STAVVNPNGNTRYFGMFKL	272		

```

RESULT 4
TN11_MOUSE
ID TN11_MOUSE STANDARD; PRM: 316 AA.
AC Q35235; Q35306; Q9RIY0; Q9JJK8; Q9JJK9;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 11 (Receptor activator
DE of nuclear factor kappa B ligand) (RANKL) (TNF-related activation-
DE induced cytokine) (TNF) (osteoprotegerin ligand) (OPGL) (osteoclast
DE differentiation factor) (ODF) (osteoclastogenesis-inhibitory factor)
DE (OCIF).
CN TNFSF11 OR RANKL OR TRANCE OR OPGL.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Hydrionema;
RX MEDLINE=97460112; PubMed=9312132;
RX Wong B.R., Rho J., Aron J., Robinson E., Orlinick J., Chao M.,
RX Karachukov S., Cayani E., Bartlett F.S. III, Frankel W.N., Lee S.Y.,
RX Choi Y.;
RA "TRANCE is a novel ligand of the tumor necrosis factor receptor family
RT that activates c-Jun N-terminal kinase in T cells.";
RL J. Biol. Chem. 272:25190-25194 (1997).
RL [2]
RP SEQUENCE FROM N.A. (ISOFORM 1).
RC TISSUE=Thymic lymphoma;
RX MEDLINE=98032977; PubMed=9367155;
RX Anderson D.M., Maraskovsky E., Billingsley W.L., Dougall W.C.,
RX Tometsko M.E., Roux E.R., Teepe M.C., Dubose R.F., Cosman D.,
RX Galibert L.;

```

RT "A homologue of the TNF receptor and its ligand enhance T-cell growth
 RT and dendritic-cell function.";
 RL Nature 390:175-179(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE-Bone marrow;
 RX MEDLINE=98227661; PubMed=9568710;
 RA Lacey D.L., Tims E., Tan H.-L., Kelley M.J., Dunstan C.R.,
 RA Burger T., Elliott R., Colombero A., Elliott G., Scully S., Hsu H.,
 RA Sullivan J., Hawkins N., Davy E., Caparrotti C., Eli A., Qian Y.-X.,
 RA Kaufman S., Sacosi I., Shalhoub V., Senaldi G., Guo J., Delaney J.,
 RA Boyle W.J.;
 RT "Osteoprotegerin ligand is a cytokine that regulates osteoclast
 RT differentiation and activation.";
 RL Cell 93:165-176(1998).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC TISSUE-Bone marrow stroma;
 RX MEDLINE=98188248; PubMed=9520411;
 RA Yasuda H., Shima N., Nakagawa N., Yamaguchi K., Kinoshita M.,
 RA Mochizuki S.-I., Tomoyasu A., Yano K., Goto K., Murakami A., Tsuda E.,
 RA Morinaga T., Higashio K., Udagawa N., Takahashi N., Suda T.;
 RT "Osteoclast differentiation factor is a ligand for
 RT osteoprotegerin/osteoclastogenesis-inhibitory factor and is identical
 RT to TRANCE/RANKL.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:3597-3602(1998).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RC STRAIN=129;
 RX MEDLINE=99214075; PubMed=10196481;
 RA Kodaira K., Kodaira K., Mizuno A., Yasuda H., Shima N., Murakami A.,
 RA Ueda M., Higashio K.;
 RT "Cloning and characterization of the gene encoding mouse osteoclast
 RT differentiation factor.";
 RL Gene 230:121-127(1999).
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORMS 1; 2 AND 3).
 RX MEDLINE=21150053; PubMed=11250921;
 RA Ikeda T., Kasai M., Utsuyama M., Hirokawa K.;
 RT "Determination of three isoforms of the receptor activator of nuclear
 RT factor-kappaB ligand and their differential expression in bone and
 RT thymus.";
 RL Endocrinology 142:1419-1426(2001).
 RN [7]
 RP SEQUENCE OF 139-147, PROCESSING, AND N-GLYCOSYLATION.
 RX PubMed=10224137;
 RA Lum L., Wong B.R., Josien R., Becherer J.D., Erdjument-Bromage H.,
 RA Schindler J., Tempst P., Choi Y., Blobel C.P.;
 RT "Evidence for a role of a tumor necrosis factor- α (TNF- α)-converting enzyme-like protease in shedding of TRANCE, a
 RT TNF family member involved in osteoclastogenesis and dendritic cell
 RT survival.";
 RL J. Biol. Chem. 274:13613-13618(1999).
 RN [8]
 RP X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 158-316.
 RX PubMed=11581298;
 RA Lam J., Nelson C.A., Ross F.P., Teitelbaum S.L., Fremont D.H.;
 RT "Crystal structure of the TRANCE/RANKL cytokine reveals determinants
 RT of receptor-ligand specificity.";
 RL J. Clin. Invest. 108:971-979(2001).
 RN [9]
 RP X-RAY CRYSTALLOGRAPHY (2.2 ANGSTROMS) OF 137-316.
 RX PubMed=11733492;
 RA Ito S., Wakabayashi K., Ubukata O., Hayashi S., Okada F., Hata T.;
 RT "Crystal structure of the extracellular domain of mouse RANK ligand at
 RT 2.2-Å resolution.";
 RL J. Biol. Chem. 277:6631-6636(2002).
 CC -1- FUNCTION: Cytokine that binds to TNFRSF1B/OPG and to
 CC TNFRSF1A/RANK. Osteoclast differentiation and activation factor.
 CC Augments the ability of dendritic cells to stimulate naive T-cell
 CC proliferation. May be an important regulator of interactions
 CC between T cells and dendritic cells and may play a role in the
 CC regulation of the T cell-dependent immune response. May also play

CC an important role in enhanced bone-resorption in humoral
 CC hypercalcemia of malignancy.
 CC -1- SUBUNIT: Homotrimer.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted
 CC (isoforms 1 and 2); Cytoplasmic (isoform 3).
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; 1 (shown here), 2 and 3; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THYMUS AND LYMPH NODES BUT
 CC NOT IN NONLYMPHOID TISSUES AND IS ABUNDANTLY EXPRESSED IN T CELLS
 CC BUT NOT IN B CELLS. A HIGH LEVEL EXPRESSION IS ALSO SEEN IN THE
 CC TRABECULAR BONE AND LUNG.
 CC -1- PTM: N-glycosylated.
 CC -1- PTM: The soluble form of isoform 1 derives from the membrane form
 CC by proteolytic processing. The cleavage may be catalyzed by
 CC ADAM17. A further shorter soluble form was observed.
 CC -1- DISEASE: DEFICIENCY IN TNFSF1 RESULTS IN FAILURE TO FORM LOBULO-
 CC ALVEOLAR MAMMARY STRUCTURES DURING PREGNANCY, RESULTING IN DEATH
 CC OF NEWBORNS. TRANCE-DEFICIENT MICE SHOW SEVERE OSTEOPEXTROSIS, WITH
 CC NO OSTEOCLASTS, MARROW SPACES, OR TOOTH ERUPTION, AND EXHIBIT
 CC PROFOUND GROWTH RETARDATION AT SEVERAL SKELETAL SITES, INCLUDING
 CC THE LIMBS, SKULL, AND VERTEBRAE AND HAVE MARKED CHONDRODYSPLASIA,
 CC WITH THICK, IRREGULAR GROWTH PLATES AND A RELATIVE INCREASE IN
 CC HYPERTROPHIC CHONDROCYTES.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF013170; AAC71061.1; -;
 DR EMBL: AF019048; AAB66812.1; -;
 DR EMBL: AF053713; AAC40113.1; -;
 DR EMBL: AB008426; BAA425425.1; -;
 DR EMBL: AB022039; BAA36970.1; -;
 DR EMBL: AB022036; BAA36970.1; JOINED.
 DR EMBL: AB022037; BAA36970.1; JOINED.
 DR EMBL: AB022038; BAA36970.1; JOINED.
 DR EMBL: AB032771; BAA97257.1; -;
 DR EMBL: AB032772; BAA97258.1; -;
 DR EMBL: AB036798; BAA97259.1; -;
 DR PDB: 1J7Z; 12-SEP-01.
 DR PDB: 1IOA; 13-MAR-02.
 DR MGI: 1100089; Tnfsl1.
 DR InterPro: IPR003636; Tnf_abc.
 DR InterPro: IPR000478; Tnf_family.
 DR Pfam: PF00229; Tnf; 1.
 DR SMART: SM00207; Tnf; 1.
 DR PROSITE: PS00251; Tnf; 1.
 DR PROSITE: PS0049; Tnf_2; 1.
 KW Cytokine; Differentiation; Receptor; Glycoprotein; Transmembrane;
 KW Signal-anchor; 3D-structure; Alternative splicing.
 FT CHAIN 1 316
 FT DOMAIN 139 316
 FT SITE 1 48
 FT TRANSMEM 49 69
 FT DOMAIN 70 316
 FT SITE 138 139
 FT CARBOHYD 197 197
 FT CARBOHYD 262 262
 FT VARSPPLIC 1 117
 FT VARSPPLIC 14 44
 FT CONFLICT 99 99 G -> D (IN REF. 2).
 FT CONFLICT 141 143 MISSING (IN REF. 5).
 SQ SEQUENCE 316 AA; 34944 MW; 08DF63A2BE00967A CRC64;

Query Match 8.8%; Score 102.5; DB 1; Length 316;
 Best Local Similarity 24.3%; Pred. No. 0.06; Mismatches 107; Indels 65; Gaps 13;
 Matches 67; Conservative 37; Mismatches 107; Indels 65; Gaps 13;

QY 3 SGLGALGILLVVSIGSWATLSAQ-EPSPQELTAEDR-----REPELNPOT 51
 D 48 SMFLALLGLGQGVCSIALFLYFRQMDPNR-----ISEDSTHCIFRILRLHFNAGLQOST 104
 QY 52 EESQDVPR-----FLEQLVRRR--SAPK-----GKARRRRAIA 84
 D 105 LESEDTLPDSCRMRKMAFOGAVQKELQHVGPFRSGAPAMMEGSLDVAQKKEPAQPF 164
 QY 85 AHYEHPPPGODGAQAGVDTVSGWEE---TKINSSPLRYDRQIGFTYIRAGLYLY 140
 D 165 AHLITNINASIPSGSH---KVTLSWYHDSGNAKISNM---LSNGKLRVNDQDFYILY 216
 QY 141 ---COVHDEGKAV---YIKLDLV---NGVALRCLSEFSATYASSPGQLRI--COV 188
 D 217 ANICFRHHETSGVPTDQLQLVVYVYKTSIKIPSSHNLMKGGSTKNMSGNSEFHFYSINV 276
 QY 189 SGLPLRFGSSLRIRTLPMWHLKAPFLTYEGLFOV 224
 D 277 GGFELRAGEEISTIOVSNPSLIDPDODATYFGAFKV 312

RESULT 5
 TNFS_CANFA
 ID TNFS_CANFA STANDARD: PRT: 260 AA.

AC 097626;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand).
 GN TNFS5 OR CD40LG OR CD40L.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hostle M.H., Willett B.J.;
 RT "Adjuvant properties of canine CD40L.";
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell
 proliferation in the absence of co-stimulus as well as IGE
 production in the presence of IL-4. Involved in Immunoglobulin
 class switching (By similarity).
 CC -1- SUBUNIT: HOMOTRIMER (By similarity).
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 EXTRACELLULAR SOLUBLE FORM (By similarity).
 CC -1- PTM: The soluble form derives from the membrane form by
 proteolytic processing (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).

CC EMBL: AF086711; AAD04375.1; -
 DR HSSP: P29965; IALY.
 DR InterPro: IPR003263; TNF_5.
 DR InterPro: IPR003636; TNF_abc.
 DR InterPro: IPR000478; TNF_FAMILY.
 DR Pfam: PF00229; TNF_1.
 DR ProDom: PD008600; TNF_5; 1.
 DR SMART: SM00207; TNF_1.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS0049; TNF_2; 1.
 CC Cytokine; Transmembrane; Glycoprotein; Signal-anchor.

FT CHAIN 1 260 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
 FT MEMBER 5, MEMBRANE FORM.
 FT CHAIN 112 260 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
 FT MEMBER 5, SOLUBLE FORM (By similarity).
 FT DOMAIN 1 22 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 23 46 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT DOMAIN 47 260 EXTRACELLULAR (POTENTIAL).
 FT SITE 111 112 CLEAVAGE (By similarity).
 FT DISULFID 177 217 POTENTIAL.
 FT CARBOHYD 239 239 N-LINKED (GLCNAC...) (POTENTIAL).
 SQ SEQUENCE 260 AA; 28688 MM; 604F69A19E9EB70 CRC64;

Query Match 7.8%; Score 90.5; DB 1; Length 260;
 Best Local Similarity 28.1%; Pred. No. 0.56; Mismatches 53; Indels 31; Gaps 5;
 Matches 39; Conservative 16; Mismatches 53; Indels 31; Gaps 5;

QY 111 ETKINSSPLRYDRQIGFT-----VIRAGLYLYLQVHFDEGKAVYIKL 155
 D 128 EASSNPASVLRWAPK-GYTTISSNVLSENGKQLAVKRQGLYVYAQYTFCSNRAASSQA 186
 QY 156 DLVNGVIALRCLSEFSAT-----AASSPGQLRLCOVS---GLPLRFGSSLRIRTL 205
 D 187 PF-----VASLCLHSPGTERVTLRAASSRGSKPCGQDSIHLGCVFELHFGASVFNVT 241
 QY 206 PMWHLKAPFLTYEGLFOV 224
 D 242 DPSOVSHGTGFTSFGLKL 260

RESULT 6
 TN11_RAT
 ID TN11_RAT STANDARD: PRT: 318 AA.

AC Q9SE2; Q91Z19;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor ligand superfamily member 11 (receptor activator
 of nuclear factor kappa B ligand) (RANKL) (TNF-related activation-
 induced cytokine) (TRANCE) (osteoprotegerin ligand) (OPGL) (osteoclast
 differentiation factor) (ODF).
 GN TNFSF11 OR RANKL OR TRANCE OR OPGL.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Ibital bone;
 RX MEDLINE=20540945; PubMed=11092398;
 RA Xu J.K., Tan J.K., Huang L., Gao X.H., Laird R., Liu D., Wysocki S.,
 RA Zheng M.H.;
 RT "Cloning, sequence and functional characterization of the rat
 homologue of receptor activator of NF-kB ligand.";
 RL J. Bone Miner. Res. 15;2178-2186(2000).
 RN [2]
 RP SEQUENCE OF 266-318 FROM N.A.
 RC STRAIN=Fischer 344;
 RX PubMed=11804028;
 RA Odgren P.R., Kim N., van Wesenbeeck L., Mackay C., Mason-Savvas A.,
 RA Satali F.F., Popoff S.N., Lengner C., van Hul W., Choi Y.,
 RA Marks S.C. Jr.;
 RT "Evidence that the rat osteopetrotic mutation toothless (tl) is not in
 the TNFSF11 (TRANCE, RANKL, ODF, OPGL) gene.";
 RL Int. J. Dev. Biol. 45:853-859(2001).
 CC -1- FUNCTION: Cytokine that binds to TNFSF11B/OPG and to
 TNFSF11A/RANK. Osteoclast differentiation and activation factor.
 CC Augments the ability of dendritic cells to stimulate native T-cell
 proliferation. May be an important regulator of interactions
 between T cells and dendritic cells and may play a role in the
 regulation of the T cell-dependent immune response. May also play
 an important role in enhanced bone-resorption in humoral
 hypercalcemia of malignancy.

```

CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -1- TISSUE SPECIFICITY: Highly expressed in thymus and bone tissues.
CC -1- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: AF187319; AAG17031.1; -
DR EMBL: AF425669; AAL23963.1; -
DR HSSP: P50591; 1D0G.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF; 1.
DR ProDom: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS00251; TNF_1; FALSE_NEG.
DR PROSITE: PS50049; TNF_2; 1.
DR Cytokine: Differentiation; Receptor; Glycoprotein; Transmembrane;
KW Signal-anchor.
FT CHAIN 1 318 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT CHAIN 141 318 MEMBER 11, MEMBRANE FORM.
FT DOMAIN 1 47 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT TRANSMEM 48 68 MEMBER 11, SOLUBLE FORM.
FT SITE 69 318 CYTOPLASMIC (POTENTIAL).
FT SITE 140 141 SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT CARBOHYD 199 264 (POTENTIAL).
FT CARBOHYD 264 264 EXTRACELLULAR (POTENTIAL).
FT CONFLICT 317 317 CLEAVAGE (BY SIMILARITY).
FT CONFLICT 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 317 317 N-LINKED (GLCNAC. . .) (POTENTIAL).
SO SEQUENCE 318 AA; 35370 MW; 4B87A4D706AD098F CRC64;

Query Match 7.8%; Score 90.5; DB 1; Length 318;
Best Local Similarity 23.3%; Pred. No. 0.72;
Matches 64; Conservative 36; Mismatches 108; Indels 67; Gaps 12;

QY 6 LALACGLLVVSLGSMATLSAO-EPSOELTAEDR-----REPELNPOTEE 54
DB 51 LALLGLGLGVVCSIALFLYFRAMDNR--ISEDSTRCFYRLRLRENTGLDSTLES 107
QY 55 QD-----VVPFLQVLVRRSA-----PKGRKARRRAIA 85
DB 108 EDTEALPDSRCRMKQAFQVAGRELOHIVGPRFSGVPAMEGSWLDVARGRPEAQPA 167
QY 86 HVEVHRPRGDDGAGVGVDSGME--TKNSSPLATYDQIGETIYRAGLYLYC 141
DB 168 HLTINADIPSGSH--KVSLSWYHDSGMAKSNMT--LSNGRLRNQOGFYLYA 219
QY 142 QV-----HFDECK--AVYLKDLLV-----NGVLALRCLLEFSATASPGQLRL--COVS 189
DB 220 NICEFHETSGSVADYDQLMAYVYVTKTSIKIPSSHNLKMGKSTKMSGNSGFHYTSINVG 279
QY 190 GLPLRPSSLRIRITLPMANILKAPFLTYGFLQV 224
DB 280 GFFKLRAGEISIVQSNPSLLDPDQATYFGARKV 314

RESULT 7
PPBN_HUMAN STANDARD; PRT; 532 AA.
AC P10696; O96CML;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Alkaline phosphatase, placental-like precursor (EC 3.1.3.1) (Nagao
DE isozyme) (germ-cell alkaline phosphatase) (PLAP-like) (ALP-1).
GN ALPL2 OR ALPL.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OC NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90283879; PubMed=2162249;
RA Lowe M.E., Straus A.W.;
RT "Expression of a Nagao-type, phosphatidylinositol-glycan anchored
RT alkaline phosphatase in human choriocarcinomas.";
RL Cancer Res. 50:3956-3962(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Choriocarcinoma;
RX MEDLINE=89308696; PubMed=2745460;
RA Watanabe S., Watanabe T., Li W.L., Soong B.-W., Chou J.Y.;
RT "Expression of the germ cell alkaline phosphatase gene in human
RT choriocarcinoma cells.";
RL J. Biol. Chem. 264:12611-12619(1989).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=88203632; PubMed=2834730;
RA Millan J.L., Manes T.;
RT "Seminoma-derived Nagao isozyme is encoded by a germ-cell alkaline
RT phosphatase gene.";
RL Proc. Natl. Acad. Sci. U.S.A. 85:3024-3028(1988).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE-Placenta;
RA Strausberg R.;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE OF 1-157 FROM N.A.
RX MEDLINE=88262578; PubMed=3387245;
RA Shen L.P., Liu H., Kan Y.-W., Kam W.;
RT "5' nucleotide sequence of a putative human placental alkaline
RT phosphatase-like gene.";
RL Nucleic Acids Res. 16:5694-5694(1988).
CC -1- CATALYTIC ACTIVITY: An orthophosphoric monoester + H(2)O = an
CC alcohol + phosphate.
CC -1- SUBUNIT: HOMODIMER.
CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a GPI-anchor.
CC -1- TISSUE SPECIFICITY: TRACE AMOUNTS IN THE TESTIS AND THYMUS,
CC AND IN ELEVATED AMOUNTS IN GERM CELL TUMORS.
CC -1- MISCELLANEOUS: IN MOST MAMMALS THERE ARE FOUR DIFFERENT ISOZYMES:
CC PLACENTAL, PLACENTAL-LIKE, INTESTINAL AND TISSUE NON-SPECIFIC
CC (LIVER/BONE/ KIDNEY).
CC -1- SIMILARITY: BELONGS TO THE ALKALINE PHOSPHATASE FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL: X55958; CAA39425.1; -
DR EMBL: J04948; AAA51700.1; -
DR EMBL: J03252; AAA98616.1; -
DR EMBL: BC014139; AAH14139.1; -
DR EMBL: X07247; CAA30232.1; AL1_SEQ.
DR PIR: A31336; A31336.
DR PIR: A34333; A34333.
DR PIR: S00974; S00974.
DR HSSP: P00634; 1ATC.
DR Stena-2DPAGE; P10696; -
DR GeneW: HGNC:441; ALPL2.
DR MIM: 171810; -

```



```

FT CARBOHYD 222 222 N-LINKED (GLCNAC..)(POTENTIAL).
FT VARSPIC 53 77 GLUTERADGAGAGGCGKPEE -> GLGFRSCORRSO
FT VARSPIC 78 244 KQISAGSOLPTS (IN ISOFORM 2).
FT VARIANT 70 70 MISSING (IN ISOFORM 2).
FT VARIANT 111 111 C -> E.
FT VARIANT 111 111 /FTID-VAR_013025.
FT CONFLICT 60 69 A -> P.
FT CONFLICT 60 69 DPGAGAOOGL -> GLSAGSGRT (IN REF. 2;
FT SEQUENCE 244 AA; 25390 MM; F41569459830E04C CRC64;
AAB37342).

Query Match 7.5%; Score 87; DB 1; Length 244;
Best Local Similarity 23.0%; Pred. No. 1.1;
Matches 64; Conservative 26; Mismatches 80; Indels 108; Gaps 13;

QY 3 SLGLAIA---CIGLLVVSILGSAWTSAGEPSELTADR-----REPEL 47
DB 18 SLILAAGATSLVTLAIPVLAVALVPQDGGVLTADPGAGAGGKPEE 77
QY 48 NPQTESQDVVPLEQLVPRRSAPKGRAPRAIAHYVHPRPGDGAQAGVDTVS 107
DB 78 EPER-----DLSPGLP-----AAHLIGAPLKGQ-----L 102
QY 108 GWERTKINS--SSPLRTDRDGEFTVIRAGLYLYCOVH-----DEKAVYKL 155
DB 103 GWETTKQAFVLTSTQTPSAEG--LALPDGLVYLCVYGRAPGSGDPGRSVYLR 161
QY 156 DLVANGVALRCLSEFSATASSPG--POLRL-----CQVS 169
DB 162 SL-----YAGAGVGPETPELLLEGAEVTPVLDPARRRGYPLMTYSVGF 208
QY 190 GLPLRPSSLRITRLPMALKAAPL--TYEGLFQV 224
DB 209 GLVLRGRERYVNI---SHPMVDFAKGTFCGAVMV 243

RESULT 9
TNF5_AOTTR
ID TNF5_AOTTR STANDARD; PRT; 261 AA.
AC Q9BDN3;
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-
DE L) (CD154 protein).
GN TNFSF5 OR CD40LG OR CD40L.
OS Aotus trivirgatus (Night monkey) (Douroucoulli).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Cebidae; Aotinae; Aotus.
OX NCBI_TaxID=9505;
RN
RP SEQUENCE FROM N.A.
RC TISSUE-Lymphocytes;
RX MEDLINE=21383618; PubMed=11491535;
RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,
RA Weiss W.R., Ansari A.A.;
RT "Cloning, sequencing, and homology analysis of nonhuman primate
RT Fas/Fas-ligand and co-stimulatory molecules.";
RL Immunogenetics 53:315-328(2001).
CC
CC -I- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell
CC proliferation in the absence of co-stimulus as well as IgE
CC production in the presence of IL-4. Involved in immunoglobulin
CC class switching (By similarity).
CC -I- SUBUNIT: HOMODIMER (By similarity).
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an
CC extracellular soluble form (By similarity).
CC -I- PTM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -I- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

```

```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL; AF344860; AAK37542.1; -.
DR HSSP; P29965; 1ALY.
DR InterPro; IPR003263; TNF_5.
DR InterPro; IPR003636; TNF_abc.
DR InterPro; IPR000478; TNF_family.
DR Pfam; PF00229; TNF_1.
DR ProDom; PD002012; TNF_abc; 1.
DR ProDom; PD008600; TNF_5; 1.
DR SMART; SM00207; TNF; 1.
DR PROSITE; PS00251; TNF_1; 1.
DR PROSITE; PS50049; TNF_2; 1.
DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 261
FT
FT CHAIN 113 261
FT
FT DOMAIN 1 22
FT TRANSMEM 23 43
FT
FT DOMAIN 44 261
FT SITE 112 113
FT DISULFID 178 218
FT CARBOHYD 240 240
FT SEQUENCE 261 AA; 29357 MM; 85EL588B507901B5 CRC64;

Query Match 7.5%; Score 87; DB 1; Length 261;
Best Local Similarity 25.9%; Pred. No. 1.2;
Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;

QY 105 TVSGWEE---TKINSSPLRYDROIGFEVIRAGLYLYCOVHPDEGKAVYKLKLLVN 160
DB 136 SVLQWAEKGYTWSNNLVTLENGKOL---TVKRGLYTYAQTVPFSNRASSAPFINS 192
QY 161 GVLAALRCLSEF-----SATASSPGPOLRLC-----QVSGLLPLRPSSLSLRITRLPM 209
DB 193 --LCKRPNNFERILLRAANTHSSAKP-----CGQGSILGIGELDPGASVFNVTDP 246
QY 210 LKAAPLTYGLFQV 224
DB 247 VSHGTGFTSGLKL 261

RESULT 10
TNF5_CALJA
ID TNF5_CALJA STANDARD; PRT; 261 AA.
AC Q9BDN3;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-
DE L) (CD154 protein).
GN TNFSF5 OR CD40LG OR CD40L.
OS Callithrix jacchus (Common marmoset).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC Callithrix.
OX NCBI_TaxID=9483;
RN
RP SEQUENCE FROM N.A.
RC TISSUE-Lymphocytes;
RX MEDLINE=21383618; PubMed=11491535;
RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,
RA Weiss W.R., Ansari A.A.;
RT "Cloning, sequencing, and homology analysis of nonhuman primate
RT Fas/Fas-ligand and co-stimulatory molecules.";
RL Immunogenetics 53:315-328(2001).
CC
CC -I- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell

```


DR InterPro: IPR003636; TNF_abc.
 DR InterPro: IPR000478; TNF_family.
 DR Pfam: PF00229; TNF_1.
 DR ProDom: PD002012; TNF_abc; 1.
 DR SMART: SM00207; TNF; 1.
 DR PROSITE: PS00251; TNF_1; FALSE_NEG.
 DR PROSITE: PS0049; TNF_2; 1.
 KW Cytochrome; Cytochrome; Transmembrane; Glycoprotein; Signal-anchor;
 KW Alternative splicing.
 FT CHAIN 1 240
 FT
 FT CHAIN 783 240
 FT
 FT DOMAIN 1 37
 FT TRANSMEM 38 58
 FT
 FT DOMAIN 59 240
 FT SITE 82 83
 FT DISULFID 154 187
 FT CARBOHYD 102 102
 FT VARSPLIC 38 73
 FT CONFLICT 120 120
 FT CONFLICT 214 214
 SO SEQUENCE 240 AA; 26351 MW; 49D0BF67E1390B39 CRC64;

Query Match 7.4%; Score 86.5; DB 1; Length 240;
 Best Local Similarity 22.2%; Pred. No. 1.2;
 Matches 57; Conservative 30; Mismatches 87; Indels 83; Gaps 10;

QY 3 SLGLALACILLVYSLG---SWATLSAQEPSELTADREPELNPQTEESDQV 58
 DB 32 SCVAVRGVGLDGLLMAGLAVOCWFLDLHWRIGEMVT---RLP-----DGPA 77
 QY 59 PLEQLVPRPRAAPKGRAPRAIAAHVEHPRGDGAAGVDTSGWEETKINSS 118
 DB 78 GSNQQLQGERSS-----HEVNPAAHLTGANSSLTG-----SGG 110
 QY 119 PLARDROI-----GEFTVIRAGLYLYLCQVHDE-----GKAVYL----- 153
 DB 111 PLMETDGLAFLRGLSYHDGALVVTAKAGYYIYSKVQLGVCGLDASTITHTGLKRT 170
 QY 154 -----KIDLVNGVLAIRCLSEFSATPAASPCQRLRCQVSGLLPLRGSLLRTPMA 208
 DB 171 PRPELELLVS-----QSPGCRATSSSRVWMDSSFLGAVVLEAGEVVRVLDER 223
 QY 209 HLKAAAP-FLTYFGLEFQV 224
 DB 224 LVRRLDGTSTSYFGAFWV 240

RESULT 12
 B3A2_MOUSE STANDARD: PRT: 1237 AA.
 ID P13808: Q9ES13; Q9ES12; Q9ES11; Q9ES10; Q9ES09;
 AC P13808: Q9ES13; Q9ES12; Q9ES11; Q9ES10; Q9ES09;
 DT 01-JAN-1990 (Rel. 13, Created)
 DT 01-JAN-1990 (Rel. 13, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Anion exchange protein 2 (Non-erythroid band 3-like protein) (B3BP).
 GN SLC4A2 OR AE2.
 OS Mus musculus (mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM A).
 RX MEDLINE=89034212; PubMed=3182834;
 RA Alper S.L., Kopito R.R., Libresco S.M., Lodish H.F.;
 RT "Cloning and characterization of a murine band 3-related cDNA from
 RT kidney and from a lymphoid cell line."
 RL J. Biol. Chem. 263:17092-17099(1988).
 RN [2]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RX MEDLINE=20462926; PubMed=11006093;

RA Lecanda J., Urtasun R., Medina J.F.;
 RT "Molecular cloning and genomic organization of the mouse AE2 anion
 RT exchanger gene."
 RT Biochem. Biophys. Res. Commun. 276:117-124(2000).
 RL -1- FUNCTION: PLASMA MEMBRANE ANION EXCHANGE PROTEIN OF WIDE
 CC DISTRIBUTION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 5 isoforms; a (shown here), b1, b2, c1 and
 CC c2; are produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: Isoform a is widely expressed at similar
 CC levels in all tissues examined. Isoforms b1 and b2 are
 CC predominantly expressed in stomach although they are also detected
 CC at lower levels in other tissues. Isoform c1 is stomach-specific.
 CC Isoform c2 is expressed at slightly higher levels in lung and
 CC stomach than in other tissues.
 CC -1- SIMILARITY: BELONGS TO THE ANION EXCHANGER FAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

CC EMBL: J04036; AA65505.1; -
 DR EMBL: AF255774; AAG23154.1; -
 DR EMBL: AF255774; AAG23155.1; -
 DR EMBL: AF255774; AAG23156.1; -
 DR EMBL: AF255774; AAG23157.1; -
 DR EMBL: AF255774; AAG23158.1; -
 DR EMBL: AF255774; AAG23157.1; -
 DR PIR: A31789; A31789.
 DR HSSP: P02730; 1B70.
 DR MGD: MGI:109351; Slc4a2.
 DR InterPro: IPR001717; Anion exchange.
 DR InterPro: IPR003020; HCO3 cotransp.
 DR Pfam: PF00955; HCO3_cotransp. 5.
 DR PRINTS: PR01231; HCO3TRANSPO.
 DR TIGRPFAMS: TIGR00834; ae; 1.
 DR PROSITE: PS00219; ANION_EXCHANGER_1; 1.
 DR PROSITE: PS00220; ANION_EXCHANGER_2; 1.
 KW Transmembrane; Glycoprotein; Transport; Antipor; Ion transport;
 KW Anion exchange; Lipoprotein; Palmitate; Alternative splicing.
 FT DOMAIN 1 703
 FT
 FT DOMAIN 704 1237
 FT TRANSMEM 704 727
 FT TRANSMEM 733 770
 FT TRANSMEM 790 812
 FT TRANSMEM 822 843
 FT TRANSMEM 844 896
 FT TRANSMEM 897 914
 FT TRANSMEM 915 929
 FT TRANSMEM 930 950
 FT TRANSMEM 984 1006
 FT TRANSMEM 1032 1053
 FT TRANSMEM 1087 1132
 FT TRANSMEM 1159 1195
 FT TRANSMEM 1195 1195
 FT DOMAIN 5 316
 FT DOMAIN 73 87
 FT DOMAIN 861 865
 FT CARBOHYD 855 855
 FT CARBOHYD 855 855
 FT CARBOHYD 866 866
 FT CARBOHYD 878 878
 FT LIPID 1169 1169
 FT VARSPLIC 1 17
 FT
 FT VARSPLIC 1 17
 FT VARSPLIC 1 166
 FT VARSPLIC 1 198
 FT VARSPLIC 167 193
 FT CONFLICT 205 205
 FT SEQUENCE 1237 AA; 136813 MW; 1A07B2C0071782EE CRC64;

Query Match	Similarity	27.9%	Score	86.5	DB	1	Length	1237
Best Local	Conservative	19	Mismatches	77	Indels	51	Gaps	12
21	GSWATLSAOPSEELT-NEDRREPELNPQTEESODVVPFLPQ----	LVRPRSAPK 74						
115	GETPTEEGEEDBESEKGRAPPO-OPSPATTSVVOFLQDEBGEKREKERSPSP 173							
75	RKAPRRRAIAAHYEHVPRPGDQAGV-----DGTSGWERTKI--N 115							
174	PTQTP-----HQEAPRASK-GAQTGLVEMVAVASATAGDDGAGRLPTKAPQH 226							
116	SSSPRYDRQIEFFVIRAGLYLKQVHFDEKAVYL--KIDL-----VNGVL-- 153							
227	RSYNLQERRRISMTGVEGA---LPRVPTDSEAOGLTATADLDMKSHRFEDVPGVRRH 283							
164	ALRCEEFSAATA---SSPGPOLR 184							
284	LVRKNAKSTQAAAREGREGPTPR 307							
RESULT 13								
TNFS_HUMAN	STANDARD:	PRT:	261	AA.				
AC	P29965:							
01-APR-1993	(Rel. 25, Created)							
01-APR-1993	(Rel. 25, Last sequence update)							
15-JUN-2002	(Rel. 41, Last annotation update)							
Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-L) (TNF-related activation protein) (TRAP) (T cell antigen gp39) (CD154 antigen)								
TNFS5 OR CD40LG OR CD40L OR TRAP.								
OS	Homo sapiens (Human).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.							
NCBI_Taxid	9606;							
SEQUENCE FROM N.A.								
MEDLINE-93076854;	Pubmed-1280226;							
Graf D., Korthauer U., Mages H.W., Briere F., Padayachee M., Malcolin S., Ugazio A.G., Notarangelo L.D., Levinsky R.J., Krocsek R.A.;								
"Cloning of TRAP, a ligand for CD40 on human T cells."								
Eur. J. Immunol. 22:3191-3194(1992).								
SEQUENCE FROM N.A.								
MEDLINE-93049181;	Pubmed-1285114;							
Hollenbaugh D., Grosmaire L.S., Kallas C.D., Chalupny J.N., Briesch-Andersen S., Noelle R.J., Stamenkovic I., Ledbetter J.A., Aruffo A.;								
"The human T cell antigen gp39, a member of the TNF gene family, is a ligand for the CD40 receptor: expression of a soluble form of gp39 with B cell co-stimulatory activity."								
EMBO J. 11:4313-4321(1992).								
SEQUENCE FROM N.A. AND VARIANTS HIGM1 128-ARG-GLY-129 AND PRO-235.								
MEDLINE-93145330;	Pubmed-7678782;							
Aruffo A., Farrington M., Hollenbaugh D., Li X., Malatovich A., Nonoyama S., Bajorth J., Grosmaire L.S., Stenkamp R., Neubauer M., Roberts R.L., Noelle R.J., Ledbetter J.A., Francke U., Ochs H.D.;								
"The CD40 ligand, gp39, is defective in activated T cells from patients with X-linked hyper-IgM syndrome."								
Cell 72:291-300(1993).								
SEQUENCE FROM N.A.								
MEDLINE-93094757;	Pubmed-1281209;							
Springer M.K., Armistage R.J., Strickbine L., Clifford K.N., Macduff B.M., Sato T.A., Maliszewski C.R., Fanslow W.C.;								
"Recombinant human CD40 ligand stimulates B cell proliferation and immunoglobulin E secretion."								
J. Exp. Med. 176:1543-1550(1992).								
SEQUENCE FROM N.A.								
MEDLINE-93138085;	Pubmed-7678552;							
Bonnefoy J.Y.;								
"Human CD40-ligand: molecular cloning, cellular distribution and regulation of expression by factors controlling IgE production."								
FEBS Lett. 315:259-266(1993).								
SEQUENCE FROM N.A.								
Shimizu M., Terasaki H., Niomiya R., Shimizu S., Nunoi H., Matsuda I.;								
Submitted (Feb-1995) to the EMBL/GenBank/DBJ databases.								
SEQUENCE OF 113-117, AND PROCESSING.								
Petravalle F., Lecocqnet-Henchoz S., Blasey H., Aubry J.-P., Elson G., Edgerton M.D., Bonnefoy J.-Y., Gauchat J.-F.;								
"Human native soluble CD40L is a biologically active trimer, processed inside microsomes."								
J. Biol. Chem. 271:5965-5967(1996).								

RA Gattl R.A., Derauf D.C., Belmont J.W., Conley M.E.;
 RT "A single strand conformation polymorphism study of CD40 ligand.
 RT Efficient mutation analysis and carrier detection for X-linked hyper
 RT Igm syndrome.";
 RL J. Clin. Invest. 97:196-201(1996).
 RN [15]
 RP VARIANTS HIGM1 ARG-36, CYS-140, SER-231: MET-254 AND GLY-227 DEL.
 RX MEDLINE=97295077; PubMed=9150729;
 RA Nonoyama S., Shimadzu M., Toru H., Seyama K., Nunoi H., Neubauer M.,
 RA Yata J.-I., Och H.D.;
 RT "Mutations of the CD40 ligand gene in 13 Japanese patients with
 RT X-linked hyper-Igm syndrome.";
 RL Hum. Genet. 99:624-627(1997).
 CC -1- FUNCTION: MEDIATES B-CELL PROLIFERATION IN THE ABSENCE OF CO-
 CC STIMULUS AS WELL AS IGE PRODUCTION IN THE PRESENCE OF IL-4.
 CC INVOLVED IN IMMUNOGLOBULIN CLASS SWITCHING.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. ALSO EXISTS AS AN
 CC EXTRACELLULAR SOLUBLE FORM.
 CC -1- TISSUE SPECIFICITY: SPECIFICALLY EXPRESSED ON ACTIVATED CD4+
 CC T-LYMPHOCYTES.
 CC -1- PTM: The soluble form derives from the membrane form by
 CC proteolytic processing.
 CC -1- DISEASE: DEFECTS IN TNFSF5 ARE THE CAUSE OF AN X-LINKED
 CC IMMUNODEFICIENCY WITH HYPER-IGM (HIGM1), AN IMMUNOGLOBULIN ISOTYPE
 CC SWITCH DEFECT CHARACTERIZED BY ELEVATED CONCENTRATIONS OF SERUM
 CC IGM AND DECREASED AMOUNTS OF ALL OTHER ISOTYPES. AFFECTED MALES
 CC PRESENT AT AN EARLY AGE (USUALLY WITHIN THE FIRST YEAR OF LIFE)
 CC RECURRENT BACTERIAL AND OPORUNISTIC INFECTIONS, INCLUDING
 CC PNEUMOCYSTIS CARINI PNEUMONIA AND INTRACTABLE DIARRHEA DUE TO
 CC CRYPTOSPORIDIUM INFECTION. DESPITE SUBSTITUTION TREATMENT WITH
 CC INTRAVENOUS IMMUNOGLOBULIN, THE OVERALL PROGNOSIS IS RATHER POOR.
 CC WITH A DEATH RATE OF ABOUT 10% BEFORE ADOLESCENCE.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -1- DATABASE: NAME=CD40Ligase;
 CC NOTE=European CD40L defect database (mutation db);
 CC WWW-http://www.expasy.org/cd40ligase/;
 CC FTP-ftp://ftp.expasy.org/databases/cd40ligase/;
 CC -1- DATABASE: NAME=PROW; NOTE=CD guide CD154 entry;
 CC WWW-http://www.ncbi.nlm.nih.gov/prov/cd/cd154.htm".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X68550; CAA48534.1; -;
 DR EMBL: Z15017; CAA78737.1; -;
 DR EMBL: X67878; CAA48077.1; -;
 DR EMBL: L07414; AAA35662.1; -;
 DR EMBL: D31797; BAA06599.1; -;
 DR EMBL: D31793; BAA06598.1; JOINED.
 DR EMBL: D31794; BAA06599.1; JOINED.
 DR EMBL: D31795; BAA06599.1; JOINED.
 DR EMBL: D31796; BAA06599.1; JOINED.
 DR PIR: S25684; S25684.
 DR PIR: S26694; S26694.
 DR PIR: S28017; S28017.
 DR PIR: S28852; S28852.
 DR PIR: JH0793; JH0793.
 DR PDB: 1ALY; 11-SEP-97.
 DR Gene: HGNC:11935; TNFSF5.
 DR MIM: 300386; -;
 DR MIM: 308230; -;
 DR InterPro: IPR003263; TNF_5.
 DR InterPro: IPR003636; TNF_abc.
 DR InterPro: IPR000478; TNF_family.
 DR Pfam: PF00229; TNF_1.
 DR ProDom: PD008600; TNF_5; 1.
 DR SMART: SM00207; TNF; 1.

DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 KM Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Antigen;
 KM Disease mutation; Polymorphism; 3D-structure.
 FT CHAIN 1 261 TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
 Query Match 7.48; Score 86; DB 1; Length 261;
 Best Local Similarity 25.9%; Pred. No. 1.4;
 Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;
 QY 105 TVSGME-----TKNSSPLRYDROIGFVIRAGLYLYCOVHFDEGKAVYIKLIDLNV 160
 DB 136 SVLQMAEKGYITWNNLVTLNGKQL--TVKRGGLYITIAQVFCNNREASSQAPPIAS 192
 QY 161 GVIALRCLEEF-----SATAASSPGPOLRLC-----QVSGLLPLRPSSLRITLPMW 209
 DB 193 --LCKSPGPFERLLRAANTHSSAKP-----CGQSTHIGCVFELQPGASVFNVDPSQ 246
 QY 210 LKAAPFLYFGFLQV 224
 DB 247 VSHGTGTFSGFLKL 261
 RESULT 14
 TNF5_MACMU
 ID TNF5_MACMU STANDARD; PRT; 261 AA.
 AC Q9BDC7;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-
 DE L) (CD154 protein).
 GN TNFSF5 OR CD40LG OR CD40L.
 OS Macaca mulatta (Rhesus macaque), and
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
 OC Cercopithecinae; Macaca.
 OX NCBI_TaxID=9544; 9531.
 RN [1]
 RP SEQUENCE FROM N.A., AND M.MULATTA VARIANTS GLN-60; VAL-204; PRO-206
 RP AND THR-215.
 RX SPECIES=M.mulatta, and C.torquatus atys; TISSUE=Lymphocytes;
 RX MEDLINE=21383618; PubMed=11491535;
 RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,
 RA Weiss W.R., Ansari A.A.;
 RT "Cloning, sequencing and homology analysis of nonhuman primate
 RT Fas/Fas-ligand and co-stimulatory molecules.";
 RL Immunogenetics 53:315-328(2001).
 CC -1- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell
 CC proliferation in the absence of co-stimulus as well as Ige
 CC production in the presence of IL-4. Involved in immunoglobulin
 CC class switching (By similarity).
 CC -1- SUBUNIT: Homotrimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an
 CC extracellular soluble form (By similarity).
 CC -1- PTM: The soluble form derives from the membrane form by
 CC proteolytic processing (By similarity).
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF344859; AAK37541.1; -;
 DR EMBL: AF344841; AAK37600.1; -;
 DR HSSP: P29965; 1ALY.
 DR InterPro: IPR003263; TNF_5.
 DR InterPro: IPR003636; TNF_abc.


```

DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR Prodom: PD002012; TNF_abc; 1.
DR Prodom: PD006600; TNF_5; 1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor; Polymorphism.
FT CHAIN 1 261
FT CHAIN 113 261
FT DOMAIN 1 22
FT TRANSMEM 23 43
FT SITE 44 261
FT DISULFID 178 218
FT CARBOHYD 240 240
FT VARIANT 60 60
FT VARIANT 204 204
FT VARIANT 206 206
FT VARIANT 215 215
SQ SEQUENCE 261 AA; 29366 MW; AA2E0F1IC0F28ABB CRC64;

Query Match
Best Local Similarity 7.4%; Score 86; DB 1; Length 261;
Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;

OY 105 TVSGWEE---TKINSSPLRYDRDIGEFTVIRAGLYLYCOVHEDEGKAVYLIKDLVLN 160
DB 136 SVLQWAEKGYTMSNNLVLENGKOL--TVKROGLYIYAQVTFCSNREASSQAPFIAS 192
OY 161 GVLALRCLEEF-----SATASSPPQRLC-----QVSGLLPLRPGSSIRIRTIPLMAH 209
DB 193 --LCUKSPGRFERILIRANHTHSAKP---CGQOSIHLAGVFELOPGASVFNVTDPSSQ 246
OY 210 LKAAPFLTYFGLFOV 224
DB 247 VSHGTGFTSFGILKL 261

RESULT 15
TNF5_MACNE STANDARD; PRT: 240 AA.
AC 09BDM7:
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DE Tumor necrosis factor ligand superfamily member 5 (CD40 ligand) (CD40-
  L) (CD154 protein).
GN TNFSF5 OR CD40LG OR CD40L.
OS Macaca nemestrina (pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9545;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lymphocytes;
RA Villinger F., Bostik P., Mayne A.E., King C.L., Genain C.P.,
  Weiss W.R., Ansari A.A.;
RT "Cloning, sequencing, and homology analysis of nonhuman primate
  Fas/Fas-ligand and co-stimulatory molecules.";
RL Immunogenetics 53:315-328(2001).
CC -i- FUNCTION: Cytokine that binds to TNFSF5. Mediates B-cell
  proliferation in the absence of co-stimulus as well as IgE
  production in the presence of IL-4. Involved in immunoglobulin
  class switching (By similarity).
CC -i- SUBUNIT: Homotrimer (By similarity).
CC -i- SUBCELLULAR LOCATION: Type II membrane protein. Also exists as an
  extracellular soluble form (By similarity).

```

```

CC -i- PRT: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -i- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AF344853; AAK37613.1; -.
CC HSSP: P29965; 1ALY.
DR InterPro: IPR003263; TNF_5.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR Prodom: PD002012; TNF_abc; 1.
DR Prodom: PD006600; TNF_5; 1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR Cytokine; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 240
FT CHAIN 113 240
FT DOMAIN 1 22
FT TRANSMEM 23 43
FT SITE 44 240
FT DISULFID 173 197
FT CARBOHYD 219 219
SQ SEQUENCE 240 AA; 27273 MW; BAE6AA0213CADEC CRC64;

Query Match
Best Local Similarity 7.3%; Score 85; DB 1; Length 240;
Matches 34; Conservative 20; Mismatches 53; Indels 24; Gaps 6;

OY 109 WEE---TKINSSPLRYDRDIGEFTVIRAGLYLYCOVHEDEGKAVYLIKDLVNGVLA 164
DB 119 MAEKGYTMSNNLVLENGKOL--TVKROGLYIYAQVTFCSNREASSQAPFIAS--LC 173
OY 165 LRCLEEF-----SATASSPPQRLC-----QVSGLLPLRPGSSIRIRTIPLMAHAKA 213
DB 174 LKSPGRFERILIRANHTHSTKP---CGQOSIHLAGVFELOPGASVFNVTDPSSVSHG 229
OY 214 PELTYFGLFOV 224
DB 230 TGFTSFGILKL 240

Search completed: March 31, 2003, 07:58:02
Job time : 18.9352 secs

```

THIS PAGE BLANK (USPTO)

RESULT 5

AB7348

alpha-glucosidase, probable [imported] - Caulobacter crescentus

C:Species: Caulobacter crescentus

C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: AB7348

R.Nierman, W.C.: Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M. Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete genome sequence of *Caulobacter crescentus*.

A:Reference number: AB7249; MUID:21173698; PMID:11259647

C:Accession: AB7348

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-670 <STO>

A:Cross-references: GB:AE005673; NID:g13422039; PIDN:AMK22781.1; GSPDB:GN00148

C:Genetics:

A:Gene: CC0796

Query Match

Best Local Similarity 7.5%; Score 87.5; DB 2; Length 670;

Best Local Similarity 24.0%; Pred. No. 10;

Matches 43; Conservative 24; Mismatches 75; Indels 37; Gaps 6;

QY 58 VPFLLEQLVPR-----RRSAPKGRKAR-----PRRAIAHYEHVPRP----- 93

DB 498 LPFTRLGLGPMIDYTPGCFRNVA PKDFKSFILPTVOTTGQALAMVYVDSPLTWSDSP 557

QY 94 -GQGAQACVD--GTVSGMEETKINSSPLRYDROIGFEVI--RAGLYLYLCOVHDE 147

DB 558 IYYAASPAOLDEISAVPTSMDEFVLSG-----EIGQYIYARKKADMMVGMATNBA 610

QY 148 GKAVYLIKLDLVNGVIALNCLFEFSATASSPGPOLRCQVSGLLPLRGSSLRIRTP 206

DB 611 GRTVAVPSFLTNGAFSAIRDEGAEPYALKTRTQSVASVSKDPLTKLAPGGGIVIRIS 669

RESULT 6

A46066

Lymphotoxin beta - human

C:Species: Homo sapiens (man)

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jul-2000

C:Accession: A46066

R.Browning, J.L.; Ngam-ek, A.; Lawton, P.; Demarinis, J.; Tizard, R.; Chow, E.P.; Hessian Cell 72, 847-856, 1993

A:Title: Lymphotoxin beta, a novel member of the TNF family that forms a heteromeric complex

A:Reference number: A46066; MUID:93208881; PMID:791655

A:Accession: A46066

A:Status: preliminary

A:Molecule type: DNA; protein

A:Cross-references: GB:LI1015; NID:g292276; PIDN:AAA36191.1; PID:g292277

A:Note: sequence extracted from NCBI Backbone (NCBI:128066, NCBI:128067)

C:Keywords: transmembrane protein

Query Match

Best Local Similarity 7.5%; Score 87; DB 2; Length 244;

Best Local Similarity 23.0%; Pred. No. 3.7;

Matches 64; Conservative 26; Mismatches 80; Indels 108; Gaps 13;

QY 3 SIGLALA---CLGLLVVSLGSMATLSAQPSEBELTAEDR-----REPEL 47

DB 18 SILLVAGATSLVTLILAVPITVLAVLPDQGLVETADPAQAQGGFLKPEE 77

QY 48 NQGTESQDVVPELQVLRPRRSAPKGRKARRRAIAHYEHVPRPGDGAQAGVDTYVS 107

DB 78 PEP-----DLSPGLP-----AAHLGAPLKGG-----L 102

QY 108 GMEETKINS--SSPLRYDROIGFEVIRAGLYLYLCOVH-----DEGKAVYIKL 155

DB 103 GMEETKEDAFILTSQGFSDABE-LALPDQGLYLYLCVLYGKRAPPGGGDPGRKSVTLRS 161

QY 156 DLVNGVIALNCLFEFSATASSPG-POLRL-----COVS 189

DB 162 SL-----YRAGAGYCPGPPELLLEGAEYTPVLDPRARQGYCPIMYTVGRC 208

QY 190 GLLPLRPGSSLRIRTLPMHLLKAAPL---TYFGLEFQV 224

DB 209 GLVQLRGRERYVNI---SHPDVDFARKGTFEGAVMV 243

RESULT 7

AB3035

Probable transmembrane sensor PA4895 [imported] - *Pseudomonas aeruginosa* (strain PA01)C:Species: *Pseudomonas aeruginosa*

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 15-Jun-2001

C:Accession: AB3035

R.Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pathogen

A:Reference number: AB2950; MUID:20437337; PMID:10984043

A:Accession: AB3035

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-340 <STO>

A:Cross-references: GB:AE004902; GB:AE004091; NID:g9951162; PIDN:AG08280.1; GSPDB:GN00148

C:Genetics:

A:Gene: PA4895

C:Superfamily: *Pseudomonas putida* regulatory protein pupR

Query Match

Best Local Similarity 7.4%; Score 86.5; DB 2; Length 340;

Best Local Similarity 24.7%; Pred. No. 6;

Matches 67; Conservative 26; Mismatches 85; Indels 93; Gaps 14;

QY 5 GLALACGLLLV-----VSLGSMATLSAQPSEBEL----- 36

DB 104 GLALLGAGPLVWGMGRQVWRDGCADYLTAVGERRDLYEDGQVEMNDSDALVDYD 163

QY 37 -----TAEDREPPELNPQTESDVVPFL-----EQLVPRRSAPKGRK 76

DB 164 AGQRLRLRYGRTYRTADNRREPR-----PFLVTRQGLMRALGTFESVR 211

QY 77 APRRAIAHYE--VHPRPGDGAQAGVDTGTVSGMEETKINSSPLRYDROIGFEVIRA 134

DB 212 BEAEVYLAIV-YEGANQVPR--EGASADGRV-----IEGQVRRDRQ-----RI 254

QY 135 GLYLYLCOVHDEKAVYIKLDLVNGVIALNCLFEFSATASSPGPOLRCQVSGLLPL 194

DB 255 G-----PVESASEALAWROGLLVADMDPLR--QWAGEIMRYGESIE-CEPS-LDPL 303

QY 195 RRGSSLRIRTLPMHLLKAAPLTYFGLEFQVH 225

DB 304 RVSGTFPVDDLP---LALAMLAQTHGLRLVH 331

RESULT 8

A31789

band 3-related protein - mouse

C:Species: Mus musculus (house mouse)

C:Date: 21-May-1990 #sequence_revision 21-May-1990 #text_change 20-Aug-1999

C:Accession: A31789

J.Alper, S.L.; Kopito, R.R.; Libresco, S.M.; Lodish, H.F. J. Biol. Chem. 263, 17092-17099, 1988

A:Title: Cloning and characterization of a murine band 3-related cDNA from kidney and liver

A:Reference number: A31789; MUID:89034212; PMID:3162834

A:Accession: A31789

A:Molecule type: mRNA

A:Residues: 1-1237 <ALP>

A:Cross-references: GB:J04036; NID:g192132; PIDN:AAA65505.1; PID:g309114

C:Superfamily: band 3 anion transport protein

C:Keywords: transmembrane protein

Query Match

Best Local Similarity 7.4%; Score 86.5; DB 2; Length 1237;

Best Local Similarity 27.9%; Pred. No. 26; Matches 57; Conservative 19; Mismatches 77; Indels 51; Gaps 12;

QY 21 GSNATLSAQPSOEELT-AEDREPELNPQTESODVVFLEQ-----LYRRSRAPK 74
 Db 115 GEMPTIEEGEEDSEASGAFRAPO-QSPATPSAVQFLQDEGAERKERTSPSP 173
 QY 75 RKAPRAIAHYEVHPRPODGAQGV-----DGTSGMEETKI---N 115
 Db 174 PTPP-----HGEAPRASK-GAQTGLVEEVAVASATAGDDGAGRPITKQPGH 226
 QY 116 SSSPLRYDQIGEFYIRAGLYLYCOVHDEGKAYL---KLDL-----VNGVL-- 163
 Db 227 RSYNLQERRRIGSMTGVEGA---LTPRPVDESEAQTATADLDMKSHRFEDVPGVRRH 283
 QY 164 ALRCLEFSATAA---SSPGPOLR 184
 Db 284 LVRRNAKSTQAAREGREGPPRR 307

RESULT 9

CD40 ligand - human
 N:Alternate names: glycoprotein 39; hCD40-L protein; T-cell antigen gp39; TRAP protein
 C:Species: Homo sapiens (man)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Apr-2001
 C:Accession: S28017; JH0793; S26694; S28852; S25684; S30593
 R:Hollenbaugh, D.; Grosmaire, L.S.; Kullas, C.D.; Chalupny, N.J.; Braesch-Anderesen, S.; EMO J. 11, 4313-4321, 1992
 A:Title: The human T cell antigen gp39, a member of the TNF gene family, is a ligand for A:Reference number: S28017; MUID:93049181; PMID:1385114
 A:Accession: S28017
 A:Molecule type: mRNA
 A:Residues: 1-261 <HOL>
 A:Cross-references: EMBL:Z15017; NID:938483; PIDN:CAA78737.1; PID:938484
 R:Spr1998, M.K.; Arnltage, R.J.; Stockbine, L.; Clifford, K.N.; Macduff, B.M.; Sato, T. J. Exp. Med. 176, 1543-1550, 1992
 A:Title: Recombinant human CD40 ligand stimulates B cell proliferation and immunoglobulin A:Reference number: JH0793; MUID:93094757; PMID:1281209
 A:Accession: JH0793
 A:Molecule type: mRNA
 A:Residues: 1-261 <SPR>
 A:Cross-references: GB:X67878; NID:938411; PIDN:CAA48077.1; PID:938412
 A:Experimental source: peripheral blood T-cell
 R:Graf, D.; Korthauer, U.; Mages, H.W.; Senger, G.; Kroccek, R.A. Eur. J. Immunol. 22, 3191-3194, 1992
 A:Title: Cloning of TRAP, a ligand for CD40 on human T cells.
 A:Reference number: S26694; MUID:93076854; PMID:1280226
 A:Accession: S26694
 A:Molecule type: mRNA
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-261 <GRA>
 A:Cross-references: EMBL:X68550; NID:937269; PIDN:CAA48554.1; PID:937270
 R:Gauchat, J.F.; Aubry, J.P.; Mazzei, G.; Life, P.; Jomotte, T.; Elson, G.; Bonnefoy, J. FEBS Lett. 315, 259-266, 1993
 A:Title: Human CD40-ligand: molecular cloning, cellular distribution and regulation of A:Reference number: S28852; MUID:93138085; PMID:7678552
 A:Accession: S28852
 A:Molecule type: mRNA
 A:Residues: 1-261 <GUA>
 A:Cross-references: EMBL:L07414; NID:9180123; PIDN:AAA35662.1; PID:9180124
 A:Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 6-Gln C:Genetics:
 A:Gene: GDB:CD40LG; HIGM1; IMD3
 A:Cross-references: GDB:120632; OMIM:308230
 A:Map position: Xq26-Xq26
 C:Keywords: glycoprotein; transmembrane protein
 F:13-44/Domain: transmembrane #status predicted <TM>
 F:45-261/Domain: extracellular #status predicted <EXT>
 F:6,240/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 7.4%; Score 86; DB 2; Length 261;
 Best Local Similarity 25.9%; Pred. No. 4.9;

Matches 35; Conservative 21; Mismatches 55; Indels 24; Gaps 6;

QY 105 TVSGWEE-----TKINSSPLRYDROIGEFYIRAGLYLYCOVHDEGKAYLKLIDLVN 160
 Db 136 SYLOWAEKGYTMSNLTLENGKOL---TVKRQGLYITTAQYTCNSNEASQAFLIS 192
 QY 161 GVALRCLEEF-----SATTAASSPGPOLRLC-----QVSGLLPLRPGSSLRIRLPMNH 209
 Db 193 --LCKLSPKRFERILLRAANTHSSAKP-----CGQGSIHGVEFELQPGASVFNVTDPQ 246
 QY 210 LKAAPLYTFGLFQV 224
 Db 247 VSHGTGTFSGFLKL 261

RESULT 10

A56764
 band 3-related protein, ileum - rabbit
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 08-Sep-1995 #sequence_revision 08-Sep-1995 #text_change 20-Aug-1999
 C:Accession: A56764
 R:Chow, A.; Dobbins, J.W.; Aronson, P.S.; Igarashi, P. Am. J. Physiol. 263, G345-G352, 1992
 A:Title: cDNA cloning and localization of a band 3-related protein from ileum.
 A:Reference number: A56764; MUID:93035730; PMID:1415547
 A:Accession: A56764
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-1237 <CHO>
 A:Cross-references: GB:S45791; NID:9256659; PIDN:AAB23488.1; PID:9256660
 A:Experimental source: New Zealand White rabbit, ileal epithelial cells
 A:Note: sequence extracted from NCBI backbone (NCBIN:115180, NCBI:115181)
 C:Superfamily: band 3 anion transport protein

Query Match 7.3%; Score 85; DB 2; Length 1237;

Best Local Similarity 28.1%; Pred. No. 35; Matches 56; Conservative 20; Mismatches 81; Indels 42; Gaps 10;

QY 21 GSNATLSAQPSOEELT-AEDREPELNPQTESODVVFLEOLYVRRSRAPKGRKAP 79
 Db 116 GATPTIEEGEEDSEANENAGARAPTEPSPAPSP-SYQFFLEDGDEGAERKERTSPSP 174
 QY 80 RRAIAHYEVHPRPODGAQGV-----DGTSGMEETKI---NSSPL 120
 Db 175 --TLPHQEAAPV-ATEGAQTGVVEEVAVASGTAGDGGASGRLTKAOPGHSYNL 231
 QY 121 RYDROIGEFYIRAGLYLYCOVHDEGKAYL---KLDL-----VNGVALRCLEE 170
 Db 232 QERRRIGSMTGAEQA---LTPRPVDESEAQTATADLDMKSHRFEDVGVRRHLYRKN 288
 QY 171 FSATAASS-----PGPOLR 184
 Db 289 AKSAGSSRREGREGPPRR 307

RESULT 11

A49266
 fas ligand - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 05-Nov-1999
 C:Accession: A49266
 R:Suda, T.; Takahashi, T.; Golstein, P.; Nagata, S. Cell 75, 1169-1178, 1993
 A:Title: Molecular cloning and expression of the Fas ligand, a novel member of the tu A:Reference number: A49266; MUID:94084792; PMID:7505205
 A:Accession: A49266
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-278 <SDU>
 A:Cross-references: GB:U03470; NID:9440178; PIDN:AAC52129.1; PID:9440179
 C:Keywords: glycoprotein; transmembrane protein

Query Match 7.2%; Score 84; DB 2; Length 278;

	Query Match	7.1%	Score 83;	DB 2;	Length 467;
	Best Local Similarity	23.6%	Pred. No. 17;		
	Matches	57;	Conservative	22;	Mismatches 91; Indels 72; Gaps 11;
OY	4 LGLALACL-----GLLLVVSLGSW--ATLSAQBSEBELTAE--DRREPP--	45			
Db	112 LGRAVGSLGRPVKRLRPLRGILRLGRAAGNRPRTUYFARGVPRIIDLHNHNRSSAETS	171			
OY	46 ---ELNPOTESODVPRLEQULV---RPRSARPKSGKAR-----PRAIQA-----	85			
Db	172 LAENIDGNPREDDLNLYRPLLGILLLRNHGKAFTTGVAUVAWDELDPGTGFPAERVALRNL	231			
OY	86 -----HYEVNRRGGOGAAGVDGVSCWEETKINSSPLRIDROIGETTIVIRAGLIY	138			
Db	232 LTGLEPPTAIRARNPFREMIGALIRADVHGW---TNQGDP-----GAAAE-----	273			
OY	139 LYCQVNFEKGAKAVUKLRDLNVGLALRCLEEFSSTAASSPGROLTCQVSGILRLRPGS	198			
Db	274 ---QHNRD-----ATLEHTANGVAHAMFAAAUVAAARRDARDVNAOLRYGSLAVVRPGS	324			
OY	199 SL 200				
Db	325 RL 326				

RESULT 15

PAHUA

alkaline phosphatase (EC 3.1.3.1) precursor, placental [validated] - human

N:Alternate names: glycerophosphatase; orthophosphoric-monoester phosphohydrolase, Regan

C:Species: Homo sapiens (man)

C:Date: 30-Sep-1997 #sequence_revision 28-Jul-1995 #text_change 08-Dec-2000

C:Accession: A1074; A27563; A24318; B24318; A25385; A94099; A94082; A28958; A90114; A24

R:Knoll, B.J.; Rothblum, K.N.; Longley, M.

J. Biol. Chem. 263, 12020-12027, 1988

A:Title: Nucleotide sequence of the human placental alkaline phosphatase gene. Evolution

A:Reference number: A1074; MUID:86296886; PMID:3042787

A:Accession: A1074

A:Molecule type: DNA

A:Residues: 1-535 <KNO>

A:Cross-references: GB:J03931; GB:M19159; NID:g178475; PIDN:AAA51710.1; PID:g178476

R:Knoll, B.J.; Rothblum, K.N.; Longley, M.

Gene 60, 267-276, 1987

A:Title: Two gene duplication events in the evolution of the human heat-stable alkaline

A:Reference number: A91586; MUID:86167830; PMID:3443302

A:Accession: A27363

A:Molecule type: mRNA

A:Residues: 1-76 <KN2>

A:Cross-references: GB:M19160

R:Henthorn, P.S.; Knoll, B.J.; Raducha, M.; Rothblum, K.N.; Slaughter, C.; Weiss, M.; La

Proc. Natl. Acad. Sci. U.S.A. 83, 5597-5601, 1986

A:Title: Products of two common alleles at the locus for human placental alkaline phosph

A:Reference number: A94113; MUID:86287303; PMID:3461452

A:Accession: A24318

A:Molecule type: mRNA

A:Residues: 23-535 <HEN>

A:Cross-references: GB:M14169

A:Note: parts of this sequence, including the amino end of the mature protein, were deter

A:Note: this is polymorphism type 1; 67-Leu also found

A:Accession: B24318

A:Molecule type: mRNA

A:Residues: 1-24, 'L', 26-65, 'V', 67-262, 'H', 264-276, 'R', 278-284, 'A', 286-388, 'C', 390-393, 'G

A:Cross-references: GB:M4170; NID:g178465; PIDN:AAA51709.1; PID:g178470

A:Note: parts of this sequence, including the amino end of the mature protein, were deter

A:Note: this is polymorphism type 3; 89-Leu also found

R:Millan, J.L.

J. Biol. Chem. 261, 3112-3115, 1986

A:Title: Molecular cloning and sequence analysis of human placental alkaline phosphatase

A:Reference number: A25385; MUID:86140079; PMID:3512548

A:Accession: A25385

A:Molecule type: mRNA

A:Residues: 1-230, 'P', 232-535 <MIL>

A:Cross-references: GB:M3077; NID:g178471; PIDN:AAC97139.1; PID:g178474

Provitl, C.E.; Straus, A.W.; Alpers, D.H.; Chou, J.Y.; Bolime, I.
Proc. Natl. Acad. Sci. U.S.A. 83, 3781-3785, 1986
A>Title: Expression of different-sized placental alkaline phosphatase mRNAs in placenta
A.Reference number: A94099; MUID:86233310; PMID:3459156
A.Accession: A94099
A:Molecule type: mRNA
A.Residues: 382-435, 'T', 437-535 <OVI>
A>Note: A soluble or serum form was also found that may arise by proteolytic cleavage
R:Kan, W.; Clausen, E.; Kim, Y.S.; Kan, Y.W.; Rutter, W.J.
Proc. Natl. Acad. Sci. U.S.A. 82, 8715-8719, 1985
A>Title: Cloning, sequencing, and chromosomal localization of human term placental alkaline phosphatase cDNA
A.Reference number: A94082; MUID:86094295; PMID:3001717
A.Accession: A94082
A:Molecule type: mRNA
A.Residues: 6-24, 'L', 26-260, 'GE', 263-323, 'H', 325-395, 'FI', 399-400, 'A', 402-535 <KAM>
A.Cross-references: GB:M1251; NID:g178463; PIDN:AA51706.1; PID:g178464
R:Milanovic, R.; Bailey, C.A.; Brink, L.; Gerber, L.; Pan, Y.C.E.; Hulmes, J.D.; Uden
Proc. Natl. Acad. Sci. U.S.A. 85, 1396-1402, 1988
A>Title: Aspartic acid-484 of nascent placental alkaline phosphatase condenses with a
A.Reference number: A28958; MUID:88144444; PMID:3422741
A.Accession: A28958
A:Molecule type: protein
A.Residues: 485-535 <MIC>
R:Exra, E.; Blacher, R.; Udenfriend, S.
Biochem. Biophys. Res. Commun. 116, 1076-1083, 1983
A>Title: Purification and partial sequencing of human placental alkaline phosphatase.
A.Reference number: A90114; MUID:84079906; PMID:6551840
A.Accession: A90114
A:Molecule type: protein
A.Residues: 23-64 <EZR>
C:Comment: this is the form of the enzyme predominantly expressed in the placenta after birth
C:Genetics:
A:Gene: GDB:ALPP; ALPI
A:Cross-references: GDB:119672; OMIM:171800
A:Map position: 2q37.1-2q37.1
A.Introns: 26/1; 65/1; 103/3; 162/1; 219/3; 264/3; 289/1; 334/1; 398/1; 437/1
A>Note: this gene is very polymorphic
C:Complex: homodimer
C:Function:
A:Description: catalyzes the hydrolysis of phosphate monoesters
C:Superfamily: alkaline phosphatase
C:Keywords: blocked carboxyl end; glycoprotein; lipoprotein; membrane protein
F:1-32/Domain: signal sequence #status predicted <SIG>
F:23-506/Product: alkaline phosphatase, placental #status experimental <MNT>
F:507-535/Domain: carboxyl-terminal propeptide #status predicted <CTP>
F:114/Active site: Ser (phosphoserine Intermediate) #status predicted
F:144,271/Binding site: carboxylate (Asn) (covalent) #status predicted
F:188/Binding site: substrate phosphate (Arg) #status predicted
F:342,382,454/Binding site: zinc (His) #status predicted
F:506/Modified site: GPR-anchor ethanolamine amidated carboxyl end (Asp) (In mature form)

```

Query Match          7.1% Score 82; DB 1; Length 535;
Best Local Similarity 27.9%; Pred. No. 24;
Matches    43; Conservative   16; Mismatches   45; Indels   50; Gaps      8

QY       76 KAPRRATAAHYEY---HPRPG---ODGAAGVDGTGSMETKINSSPLRYDRDIGEF 129
        ||| :|| |         || :||| :|| :|| :|| :|| :||
DB       403 KARDKRA-----YTLLLYGGNGPGVLTGDKARPDVTESSGSPXYRGOASVPDLEETHAGED 458
QY       130 TYIRAGLYLTYCQHFDDEKKAVYLKLIDLNVN-----LALRCLEESA----- 173
                || :|| :|| :|| :|| :|| :|| :|| :||
DB       459 VAV-----FARGPOAH-----LVHGQEDOTFIAYMAFAACLEPYTADCLAPP 501
QY       174 ---TAASSPGLRLCOVSGSLPLRPBGSSLRIPT 204
        || :|| |         ||| ||| :|| :|| :|| :||
DB       502 AGTTDAHPGRSV-----YPALLPLLAGTLLET 531

Search completed: March 31, 2003, 07:59:15
Job time : 33.5442 secs
```

Search completed: March 31, 2003, 07:59:15
Job time : 33.5442 secs

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 00:47:43 ; Search time 232.957 Seconds
(without alignments)
2175.080 Million cell updates/sec

Title: US-09-245-198A-2

Perfect score: 1162

Sequence: 1 VLSGLALACILGLLVVSL.....PMAHLKAPFLTYGFLGVH 225

Scoring table:

BLOSUM62	
Xgapop 10.0 , Xgapext 0.5	
Ygapop 10.0 , Ygapext 0.5	
Fgapop 6.0 , Fgapext 7.0	
Delop 6.0 , Delext 7.0	

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

```

-MODEL=frame+ p2n.model -DEV=xlh
-O=/cgn2.1/USPO.spool/US09245198/unat_24032003.163552.25944/app.query.fasta.1.846
-DB=N_Geneseq_101002 -QFMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09245198 -QGN_1.1.338 -tunat_24032003.163552.25944 -NCPU=6 -ICPU=3
-NO_XLPRX -NO_MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

```

Database :

```

1: N_Geneseq_101002.*
2: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1980.DAT.*
3: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1981.DAT.*
4: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1982.DAT.*
5: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1983.DAT.*
6: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1984.DAT.*
7: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1985.DAT.*
8: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1986.DAT.*
9: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1987.DAT.*
10: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1988.DAT.*
11: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1989.DAT.*
12: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1990.DAT.*
13: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1991.DAT.*
14: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1992.DAT.*
15: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1993.DAT.*
16: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1994.DAT.*
17: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1995.DAT.*
18: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1996.DAT.*
19: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1997.DAT.*
20: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1998.DAT.*
21: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA1999.DAT.*
22: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2000.DAT.*
23: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001A.DAT.*
24: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2001B.DAT.*
25: /SIDS2/gcgcdata/geneseq/geneseqn-emb1/NA2002.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1162	100.0	1168	19	AAV18599	Mus musculus tumor
2	1089	93.7	701	20	AAV23425	Mus musculus tumor
3	1020	87.8	1236	19	AAV47613	TNF related endoth
4	1020	87.8	1236	22	AAV04350	Human TREPA (TNF r
5	1020	87.8	1353	21	AAV49717	Human PRO207 CDNA
6	1020	87.8	1353	24	ABK40255	CDNA encoding huma
7	1020	87.8	1353	24	ABK40255	CDNA encoding huma
8	1020	87.8	1373	19	AAV18600	Homo sapiens tumor
9	1020	87.8	1421	20	AAV56000	Human tumor necro
10	951	81.8	898	22	AAV03964	Expression vector
11	951	81.8	1030	20	AAV23424	Human TNRL3 DNA.
12	113.5	9.8	2148	24	ABK11680	DNA encoding tumor
13	111	9.6	1221	24	ABK11679	DNA encoding tumor
14	110.5	9.5	978	23	ABL21473	Drosophila melanog
15	106.5	9.2	1630	19	AAV41377	NF-kB receptor act
16	106.5	9.2	1630	19	AAV41377	NF-kB receptor act
17	106.5	9.2	1630	22	AAV15310	Murine receptor ac
18	106.5	9.2	1630	22	AAV08714	Murine receptor ac
19	106.5	9.2	1630	22	AAV05903	Human secreted pro
20	106.5	9.2	3878	22	AAV05230	Human prosate CDNA
21	106.5	9.2	4034	22	AAV64040	Human prosate CDNA
22	106.5	9.2	4034	22	AAV93868	Human prosate CDNA
23	106.5	9.2	4034	22	ABL95411	Human prosate CDNA
24	106.5	9.2	4894	22	AAV64038	Human prosate CDNA
25	106.5	9.2	4894	22	AAV93866	Human prosate CDNA
26	106.5	9.2	4894	22	ABL95409	Human prosate CDNA
27	106.5	9.2	6976	22	AAV64041	Human prosate CDNA
28	106.5	9.2	6976	22	AAV93869	Human prosate CDNA
29	106.5	9.2	6976	22	ABL95412	Human prosate CDNA
30	105.5	9.1	81800	24	ABK84756	Human prosate CDNA
31	103	8.9	6804	22	AAV82935	Human prosate CDNA
32	102.5	8.8	951	19	AAV69900	Nucleotide sequenc
33	102.5	8.8	951	21	AAV39156	Mouse OSM nucleoti
34	102.5	8.8	951	21	AAV29965	DNA encoding a mur
35	102.5	8.8	951	21	AAV49024	Osteoclast formati
36	102.5	8.8	1538	19	AAV69886	Nucleic acid encod
37	102.5	8.8	2166	24	ABK11678	DNA encoding tumor
38	102.5	8.8	2191	19	AAV41489	Nucleotide sequenc
39	102.5	8.8	2237	20	AAV80224	Murine TRANCE enco
40	102.5	8.8	2237	24	ABK12880	CDNA encoding mus
41	102.5	8.8	2295	19	AAV70284	Human osteoprotege
42	102.5	8.8	2299	21	AAV99966	DNA encoding a mur
43	101.5	8.7	7332	24	ABL40768	Constitutively act
44	99.5	8.6	38186	20	AAV32028	Human METH1 relate
45	99.5	8.6	38186	22	AAV90085	AC004449 CDNA clon

ALIGNMENTS

RESULT 1
AAV18599 standard; CDNA, 1168 BP.

AAV18599;
21-JUL-1998 (first entry)

Mus musculus tumor necrosis factor related ligand (TRELL) gene.

TRELL: tumor necrosis factor related ligand; tnfr; treatment;
cancer; autoimmune disease; immune system; stimulation; suppression;
graft rejection; ds.

Mus musculus.

Key Location/Qualifiers
CDS 2..679

```

FT      /tag- a
FT      /note= "tumour necrosis factor related ligand"
XX
XX
XX      MO9805783-A1.
XX
XX      12-FEB-1998.
XX
XX      07-AUG-1997: 97WO-US133945.
XX
XX      18-MAR-1997: 97US-0040820.
XX      07-AUG-1996: 96US-0023541.
XX      18-OCT-1996: 96US-0028515.
XX
XX      (BIO ) BIOGEN INC.
XX      (UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX
XX      Browning JL, Chicheportliche Y;
XX
XX      WPI: 1998-145619/13.
XX      P-PSDB: AAM47524.
XX
XX      Tumour necrosis factor related ligand - useful for, e.g. treating
XX      cancer, auto-immune disease and immune responses to tissue grafts
XX
XX      Claim 2; Pages 45-46; 69pp; English.
XX
XX      The sequence is that encoding mouse tumour necrosis factor related
XX      ligand (TRELL). TRELL or active fragments can be included with a
XX      carrier in pharmaceutical compositions to treat cancer, auto-immune
XX      diseases or immune responses to tissue grafts, or to stimulate or
XX      suppress the immune system. It is useful to screen for TRELL
XX      receptors, by labelling with a detectable label and screening
XX      compositions for binding. Agents interfering with TRELL-receptor
XX      binding can also be screened for, can then be administered,
XX      optionally with interferon- gamma, to induce cell death or
XX      treat, suppress or alter immune responses (especially involving human
XX      adenocarcinoma cells) involving a signal pathway between TRELL and its
XX      receptor. The DNA sequence can be used in gene therapy for
XX      TRELL-related disorders in mammals (especially humans), e.g. tumours,
XX      auto-immune and inflammatory diseases or inherited genetic disorders,
XX      by introducing into cells, and expressing, therapeutically effective
XX      amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
XX      It may also be of use in the preparation of prepare probes for
XX      screening natural/synthetic DNAs for TRELL-encoding sequences
XX      and for antisense therapy.
XX
XX      Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;
XX
XX      Alignment Scores:
XX      Pred. No.: 8.02e-97 Length: 1168
XX      Score: 1162.00 Matches: 225
XX      Percent Similarity: 100.00% Conservative: 0
XX      Best Local Similarity: 100.00% Mismatches: 0
XX      Query Match: 100.00% Indels: 0
XX      DB: 19 Gaps: 0
XX
XX      US-09-245-198a-2 (1-225) x AAV18599 (1-1168)
XX
XX      1 ValIeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValIserLeu 20
XX      2 GTCTGAGGCTGGGCGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG 61
XX
XX      21 GlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAsp 40
XX      62 GGGAGCTGGGCAACGCTGTGGCCAGAGACCTTCTCAGAGAGAGCGAGACGAGAGAC 121
XX
XX      41 ArgArgGluProGluLeuAsnProGluThrGluGluSerGlnAspValValProPhe 60
XX      122 CGCGGGAGAGCCCTGAACTGAATCCCAAGAGAGAAAGCAGATGTGTACCTTTC 181
XX
XX      61 LeuGluGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArg 80
XX      182 TTGGAACTACTAGTCCGGCTCGAAGAAAGTCTCTAAAGGCGGAGAGCGGCGCTCGC 241

```

```

XX      81 ArgAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 100
XX      242 CGAGCTATTGACAGCCCATTTATGAGTTCACTCCCGGCGAGAGATGAGACAGCAAGCA 301
XX
XX      101 GlyValAspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerProLeu 120
XX      302 GGTGTGATGAGGACAGTGAAGTGGCTGGGAAAGCAACCAATCAACAGCTTCACCCCTG 361
XX
XX      121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTrpLeuTyr 140
XX      362 CGCTACGACCCGACAGATGGGGAATTTACAGTCATAGGCTGGCTGCTACTACCTGTAC 421
XX
XX      141 CysGlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuValAsn 160
XX      422 TGTACAGTGCACCTTGATGAGGAAAGCGCTGTACTAGCTGAGACTTGGTGGTGAAC 481
XX
XX      161 GlyValLeuAlaLeuValArgCysLeuGluGluPheSerAlaThrAlaAlaSerProGly 180
XX      482 GGTGTGCTGGCCCTGGGCTGCTGGAAATTCACAGCCACAGCAGCAAGCTTCCTGGG 541
XX
XX      181 ProGlnLeuValArgCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeu 200
XX      542 CCCAGCTCCGTTGTGCGCAGGTGTCTGGCTGTGCGCTGGCGGCGCAGGGCTTCCCTT 601
XX
XX      201 ArgIleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGly 220
XX      602 CGGATCGGACCCCTCCCTGGCTCAATCTTAAGCTGTGCCCCCTTCAACCTACTTGGGA 661
XX
XX      221 LeuPheGlnValHis 225
XX      662 CTCCTTCAAGTTCAC 676
XX
XX      RESULT 2
XX      AAX23425
XX      AAX23425 standard; DNA: 701 BP.
XX
XX      AAX23425;
XX
XX      18-JUN-1999 (first entry)
XX
XX      Mouse TNRL3 DNA.
XX
XX      Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
XX      developmental abnormality; gestational abnormality; prostate cancer;
XX      APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
XX      cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
XX      apoptosis; mouse; ss.
XX
XX      Mus sp.
XX
XX      Key 1.636 Location/Qualifiers
XX      CDS 1..636 /tag- a
XX      FT /product= "TNRL3"
XX
XX      MO9911791-A2.
XX
XX      11-MAR-1999.
XX
XX      04-SEP-1998: 98WO-US18393.
XX
XX      05-SEP-1997: 97US-0924634.
XX
XX      (UNIV ) UNIV WASHINGTON.
XX
XX      Chaudhary PM;
XX
XX      WPI: 1999-205191/17.
XX      P-PSDB: AAM93591.
XX
XX      New Tumour Necrosis Factor family receptor polypeptides and ligands -
XX      useful for diagnosis and treatment of prostate cancer and

```

PT developmental or gestational abnormalities

XX Example VII; Fig 13B; 156pp; English.

XX This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: ApO4, ApO6, ApO8 and ApO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. ApO4 is useful for diagnosing prostate cancer
 CC by determining levels of ApO4 in an individual. Prostate cancer can also
 CC be treated using ApO4 selective binding agents linked to a therapeutic
 CC moiety. ApO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. ApO4 polypeptides/active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in ApO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using ApO4 polypeptides/active
 CC fragments and ApO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of ApO4 and detecting a change in level of ApO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. ApO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. ApO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.

XX Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other:

Alignment Scores:

Pred. No.: 1,94e-90 Length: 701
 Score: 1089.00 Matches: 210
 Percent Similarity: 99.53% Conservative: 0
 Best Local Similarity: 99.53% Mismatches: 1
 Query Match: 93.72% Indels: 0
 DB: Gaps: 0

US-09-245-198A-2 (1-225) x AAX23425 (1-701)

OY 15 LeuValValSerLeuGlySerTyrPalThrLeuSerAlaGlnGluProSerGlnIleu 34
 DB 1 CTGGCTGGTGGAGCTGGGAGCGGCAACGCTGTGCCAGAGACCTTCTAGAG 60
 OY 35 GluLeuThrAlaGluAspArgGluProGluLeuAsnProGlnThrGluIleuSer 54
 DB 61 GAGCTGAG 120
 OY 55 GlnAspValValProPheLeuGlnGlnLeuValArgProArgGlySerAlaProIleu 74
 DB 121 CAGGATGTGTACCTTCTTGGACAACTAGTCCGCGCTCGAAGAGTCTCTAAAGC 180
 OY 75 ArgLysAlaArgProArgAlaAlaIleAlaHisTyrGlnValHisProArgProGly 94
 DB 181 CGGAAGGGGGGCGCTCGCGAGCTATTCACGCCATTTAGAGTTTCACTCGGCCAGGA 240
 OY 95 GlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrGlnGluThrLysIle 114
 DB 241 CAGATGGAGCACAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATC 300
 OY 115 AsnSerSerSerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIleArgAla 134
 DB 301 AACAGCTCCAGCCCTCTCGCTACGACCGCAGATTGGGAAATTAAAGTATCAGGGCT 360
 OY 135 GlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlyGlyLysAlaValTyrIleuLys 154
 DB 361 GGGCTCTACTACTGTACTGTGAGGTGACATTGTGATGAGGGAAGGCTGTACTCTAGAG 420
 OY 155 LeuAspLeuLeuValAsnGlyValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThr 174
 DB 421 CTGGACTTGGCTGGTAAGAGGTGTCTGCGCCCTCGCTGGAGAAATTCACACCACA 480
 OY 175 AlaAlaSerSerProGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeu 194
 DB 481 GCAGCAAGCTCTCTGCGGCCAGCTCCGTTGTGTCCAGAGTGTGTGGCTGTGGCGCTG 540

OY 195 ArgProGlySerSerLeuArgIleArgThrLeuProThrAlaHisLeuLysAlaAlaPro 214
 DB 541 CGGCCAGGGCTCTCCCTTCGGATCCGACACCTCCCTCGGCTCATCTTAAGGCTGCCCC 600

OY 215 PheLeuThrTyrPheGlyLeuPheGlnValHis 225
 DB 601 TTCCTAACCTACTTTGGACTCTTCAAGTCAAC 633

RESULT 3

AAV47613
 ID AAV47613 standard; cDNA; 1236 BP.

AAV47613;

27-OCT-1998 (first entry)

TNF related endothelium proliferative agent gene.

ss; TNF; endothelium proliferative agent; TREPA; wound healing; cancer;

tissue grafting; vascularisation; apoptosis; autoimmune; birth control.

Homo sapiens.

Key Location/Qualifiers

CDS 1..750
 /*tag= a
 /product= "TREPA"

W09835061-A2.

13-AUG-1998.

12-FEB-1998; 98WO-US02859.

10-FEB-1998; 98US-0021706.

12-FEB-1997; 97US-0798692.

(ABBO) ABBOTT LAB.

Wiley SR;

WPI: 1998-447255/38.

P-PSDB; AAM29745.

Detecting nucleic acid encoding TREPA - useful for diagnosis and

treatment of autoimmune disease, tumours and inflammation

Claim 11; Page 123-4; 142pp; English.

The TNF-related endothelium proliferative agent (TREPA), or its

activators or agonists, are used to treat a deficit of TREPA, e.g. to

promote wound healing or tissue grafting, by promoting vascularisation,

also to induce apoptosis for treating cancer and eliminating autoreactive

T cells, as an adjunct to cancer chemotherapy or antiviral treatment.

TREPA peptides can also be used to target cytotoxic agents or for

affinity isolation of the corresponding receptor, the nucleic acid for

which can be used to transform tumour cells to render them more

responsive to TREPA and to screen for TREPA mimics.

Ribozymes, antisense RNA, antibodies or peptides, are used to treat

TREPA-associated diseases, e.g. tumours and metastases (by inhibiting

vascularisation), inflammation or a wide range of autoimmune conditions,

conditions involving abnormal stimulation of epithelial cells (e.g.

atherosclerosis), for birth control (inhibiting ovulation and placental

formation) or other angiogenic conditions (e.g. ulcers).

Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:

Alignment Scores: 7.98e-84 Length: 1236
 Pred. No.: 1020.00 Matches: 199
 Percent Similarity: 92.86% Conservative: 9
 Best Local Similarity: 88.84% Mismatches: 16

[illegible][illegible]

```

Oy 142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
Db 496 CAGGTGCACCTTGTGAGGGGAGGCTGTCTACCTGAAGCTGACCTGTGGTGGATGGT 555
Oy 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181
Db 556 GTGCGGGCCCTGGCGTGTGCTGAGGAATTCTCAGCCACTGGCGGAGATTCTCCCTGGGCC 615
Oy 182 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
Db 616 CAGCTCCCGCTCTGTGCAGGTGTCTGGCTGTGGCCCTGGCGGACAGGCTCTCCCTGGCG 675
Oy 202 IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
Db 676 ATCCGACCCCTCCCTGGGCCCCATCTCAAGGCTGCCCTTCCTCCTACTTCGGACTTC 735
Oy 222 PheGlnValHis 225
Db 736 TTCGAGGTTCAC 747

RESULT 5
AAA49717
ID AAA49717 standard; cDNA; 1353 BP.
AC AAA49717;
XX 25-SEP-2000 (first entry)
DT
XX Human PRO207 cDNA clone DNA30879-1152.
DE
XX PRO207; human; antitumour; tumour; therapy; cytostatic;
KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.
XX Homo sapiens.
OS
XX Key Location/Qualifiers
FH CDS 58..807
FT sig_peptide 58..177 /*tag= a
FT mat_peptide 178..804 /*tag= b
FT /*tag= c
XX WO200037638-A2.
XX 29-JUN-2000.
PD
XX 02-DEC-1999; 99WO-US28565.
PF
XX 22-DEC-1998; 98US-0113296.
PR 08-MAR-1999; 99WO-US05028.
PR 21-APR-1999; 99US-0130232.
PR 28-APR-1999; 99US-0131445.
PR 14-MAY-1999; 99US-0134287.
PR 20-JUL-1999; 99US-0144758.
PR 26-JUL-1999; 99US-0145698.
PR 15-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;
XX Napier MA, Pitti RM, Wood WI;
XX WPI: 2000-442668/38.
XX P-PSDB: AAY95338.
XX
XX Novel composition to inhibit neoplastic cell growth or for treating
XX tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,
XX PRO221, PRO324, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or

```

```

PT PRO666 -
XX
XX Claim 20; Fig 3; 172pp; English.
PS
XX
XX The present sequence is that of cDNA clone DNA30879-1152
CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows
CC homology to several members of the tumour necrosis factor family,
CC especially human lymphotoxin (23.4%). The cDNA was identified in a
CC foetal kidney cDNA library following identification of an expressed
CC sequence tag with homology to human Apo-2 ligand. A claimed method
CC for inhibiting the growth of a tumour cell comprises exposing the
CC tumor cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,
CC PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO666 (see
CC AAY95337-49), their agonists or chimeric polypeptides incorporating
CC them. The tumour is especially a cancer selected from breast,
CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
CC central nervous system cancer, melanoma and leukaemia. Nucleic
CC acids encoding PRO179 etc. are used in the recombinant production
CC of the antitumour polypeptides.
XX
XX Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;
SQ
Alignment Scores:
Pred. No.: 8.95e-84 Length: 1353
Score: 1020.00 Matches: 199
Percent Similarity: 92.86% Conservative: 9
Best Local Similarity: 88.84% Mismatches: 16
Query Match: 87.78% Indels: 0
DB: 21 Gaps: 0
US-09-245-198a-2 (1-225) x AAA49717 (1-1353)
Oy 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
Db 133 CTGGGCTGGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCG 192
Oy 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41
Db 193 AGCGGGGATGCGTGTGCTGCCAGAGCTCCAGAGAGAGAGTGTGTGAGAGAGAGAC 252
Oy 42 ArgGluProProGluLeuAsnProGlnThrGluSerGlnAspValValProPheLeu 61
Db 253 CAGGACCGCTGGGAATCTGAATCCCGACAGAAACAGAGATCCCTGCTCTGTG 312
Oy 62 GlnLeuLeuValArgProArgSerAlaProLysGlyValArgValAlaArgArg 81
Db 313 AACGACTAGTTCGGCTCGCAGAGAGTCCACCTTAAGCCCGGAACACGGGCTCGAAGA 372
Oy 82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
Db 373 GCGATCGCAGCCCATTTATGAAGTTTCATCCAGCACTGACAGAGAGAGCGGACGAGGT 432
Oy 102 ValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerProLeuArg 121
Db 433 GTGGACGGGACAGTGAAGTGTGGCTGGAGGAGACCAAGAACTCAAGCTCCAGCCCTTGGCG 492
Oy 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLys 141
Db 493 TACACCGCCCGACATGGGAGATTATGATCAACCGGCTGGGCTCTACTACTGACTGCT 552
Oy 142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
Db 553 CAGGTGCACCTTGTGAGGGGAGGCTGTCTACCTGAAGCTGACCTGTGGTGGATGGT 612
Oy 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181
Db 613 GTGCGGGCCCTGGCGTGTGCTGAGGAATTCTCAGCCACTGGCGGAGATTCTCCCTGGGCC 672
Oy 182 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
Db 673 CAGCTCCCGCTCTGTGCAGGTGTCTGGCTGTGGCCCTGGCGGACAGGCTCTCCCTGGCG 732
Oy 202 IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221

```

DB 733 ATCCGACCCCTCCCGGGCCCATCTCAAGCGTCCCTTCTCCTACTTCGAGCTC 792
 OY 222 PheGlnValHis 225
 DB 793 TTCACGCTTCAC 804
 RESULT 6
 ABR40255
 ID ABR40255 standard; cDNA; 1353 BP.
 XX ABR40255;
 AC ABR40255;
 DT 15-JUL-2002 (first entry)
 XX
 DE cDNA encoding human PRO207 polypeptide.
 XX
 KW Human: PRO; benign tumour; malignant tumour; lymphoid malignancy;
 KW Leukemia; neuronal disorder; stromal disorder; blastocoealic disorder;
 KW Inflammatory disorder; immune disorder; angiogenic disorder;
 KW gene therapy; cytostatic; neuroprotective; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200153486-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 11-FEB-2000: 2000MO-US03565.
 XX
 PR 08-MAR-1999: 99MO-US05028.
 PR 11-MAR-1999: 99US-123972P.
 PR 11-MAY-1999: 99US-133459P.
 PR 02-JUN-1999: 99MO-US12252.
 PR 22-JUN-1999: 99US-140650P.
 PR 22-JUN-1999: 99US-140653P.
 PR 20-JUL-1999: 99US-144758P.
 PR 26-JUL-1999: 99US-145698P.
 PR 28-JUL-1999: 99US-146222P.
 PR 17-AUG-1999: 99US-149395P.
 PR 31-AUG-1999: 99US-151689P.
 PR 01-SEP-1999: 99MO-US20111.
 PR 15-SEP-1999: 99MO-US21090.
 PR 30-NOV-1999: 99MO-US28313.
 PR 01-DEC-1999: 99MO-US28301.
 PR 01-DEC-1999: 99MO-US28634.
 PR 05-JAN-2000: 2000MO-US00219.
 XX
 XX (GENE) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
 PI Marsters SA, Pan J, Pletli RM, Roy MA, Smith V, Stone DM;
 PI Watanabe CK, Wood WI.
 XX
 DR WPI: 2002-205567/26.
 DR P-PSDB: AA086129.
 XX
 XX Thirty five nucleic acids encoding PRO polypeptides, useful for
 PT treating benign or malignant tumours, leukemias and lymphoid
 PT malignancies, inflammatory, angiogenic and immunologic disorders -
 XX
 PS Claim 50: Fig 3: 302pp: English.
 XX
 XX The present invention relates to the isolation of novel human PRO
 CC polypeptides and the polynucleotide sequences encoding them. The
 CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
 CC useful for treating benign or malignant tumours (e.g. renal, kidney,
 CC bladder, breast, etc), leukemias and lymphoid malignancies, other
 CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
 CC macrophagal, stromal and blastocoealic disorders, inflammatory, immune
 CC and angiogenic disorders. The polynucleotide sequences are also
 CC useful in gene therapy. ABR40254-ABR40288 encode for the human PRO
 CC polypeptides of the invention.

XX
 SQ Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;
 Alignment Scores:
 Pred. No.: 8 95e-84 Length: 1353
 Score: 1020.00 Matches: 199
 Percent Similarity: 92.86% Conservative: 9
 Best Local Similarity: 88.84% Mismatches: 16
 Query Match: 87.78% Indels: 0
 DB: 24 Gaps: 0
 US-09-245-198A-2 (1-225) x ABR40255 (1-1353)
 OY 2 LeuSerLeuGlyLeuValLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
 DB 133 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 192
 OY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41
 DB 193 AGCGGCGCATCGCTGTCGCGCCAGAGCGCTGCGCCAGAGAGAGCGTGGCAGAGAGAG 252
 OY 42 ArgGluProProGluLeuAsnProGluThrGluSerGlnAspValValProPheLeu 61
 DB 253 CAGGACCCGTCGGAAGTGAATCCCAAGACAGAAAGACAGAGATCTCGGCTTCTCTG 312
 OY 62 GlnGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArgArg 81
 DB 313 AACCGACTAGTTGGGCTGCGCCAGACAGTGCACCTAAAGGCGGAAACAGGGCTCGAGA 372
 OY 82 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspLysAlaGlnAlaGly 101
 DB 373 GCGATCGACAGCCCATATGATGATCCACGACCTGACAGACGACGACGACGACGACG 432
 OY 102 ValAspGlyThrValSerGlyTyrGluGluThrLysIleAsnSerSerProLeuArg 121
 DB 433 GTGACGGGACAGTGAATGCTGGAGAGACCAAGATCAAGCTTCAGCCCTCTGCGC 492
 OY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
 DB 493 TACACCGCCAGATCGGGAGTTATGATGACACCGGGGCGTGTACTGACTGACTG 552
 OY 142 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
 DB 553 CAGGTTCACACTTATGATGAGGAGGAGGCTGTACTGACTGACTGACTGACTGACTG 612
 OY 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSerSerProGlyPro 181
 DB 613 GTGCTGGCGCTGGCTGCTGGAGGAATTTCTACGCCATCGCGGAGTTCCCTCGGGCCC 672
 OY 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
 DB 673 CAGCTCGGCTGTGCGAGGAGTGTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 732
 OY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
 DB 733 ATCCGACCCCTCCCGGGCCCATCTCAAGCGTCCCTTCTCCTACTTCGAGCTC 792
 OY 222 PheGlnValHis 225
 DB 793 TTCACGCTTCAC 804
 RESULT 7
 ABR34881
 ID ABR34881 standard; cDNA; 1364 BP.
 XX
 AC ABR34881;
 XX
 DT 08-MAY-2002 (first entry)
 XX
 DE Human cDNA encoding secreted protein #19.
 XX
 KW Human: secreted protein; gene; ss; nutritional supplement; haemophilic;
 KW viral infection; bacterial infection; fungal infection; diabetes; asthma;

KW autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumour;
 KW autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
 KW Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
 KW coagulation disorder; inflammatory disorder; Crohn's disease; incision;
 KW tissue regeneration; wound healing; burn; haematopoiesis;
 KW myeloid cell deficiency; lymphoid cell deficiency.
 OS
 XX Homo sapiens.
 XX
 XX MO200177288-A2.
 XX
 XX 18-OCT-2001.
 XX
 XX 29-MAR-2001; 2001WO-US10224.
 XX
 XX 06-APR-2000; 2000US-195582P.
 XX
 XX (GBV) GENETICS INST INC.
 XX
 XX Mong GG, Clark HF, Fachtel K, Agostino MJ, Howes SH, Resnick RJ;
 PI Guluketa K, Graham JR;
 XX
 XX WPI; 2002-179321/23.
 XX
 XX Five hundred and ninety two polynucleotides derived from a variety of
 PT human tissue sources which encode secreted proteins, useful for
 PT treating immune deficiencies and disorders such as autoimmune disorders
 PT
 XX
 XX Claim 1; Page 82; 372pp; English.
 XX
 XX The invention relates to 592 polynucleotides which have been derived from
 CC a variety of human tissue sources and which encode novel secreted
 CC proteins. The polynucleotides can be used as probes for the
 CC identification and isolation of full length cDNA and genomic DNA. The
 CC polynucleotides and proteins can also be used as nutritional supplements.
 CC The proteins are useful in the treatment of various immune deficiencies
 CC and disorders such as viral infections, bacterial infections, fungal
 CC infections, autoimmune disorders (e.g. rheumatoid arthritis, multiple
 CC sclerosis, autoimmune thyroiditis and diabetes) and allergic reactions
 CC and conditions (e.g. asthma). They are also useful for treating
 CC neurodegenerative diseases (e.g. Alzheimer's disease, Parkinson's
 CC disease), liver fibrosis, coagulation disorders (e.g. haemophilia),
 CC inflammatory disorders (e.g. Crohn's disease) and tumours. They are also
 CC useful for tissue regeneration, for wound healing and in the treatment of
 CC burns, incisions and ulcers. The proteins are also useful for regulating
 CC haematopoiesis and for treating myeloid or lymphoid cell deficiencies.
 CC Sequences ABR34863-ABR35454 represent polynucleotides of the invention.
 XX
 XX S0 Sequence 1364 BP; 246 A; 461 C; 394 G; 263 T; 0 other;
 XX
 XX Alignment Scores:
 XX Pred. No.: 9.04e-84 Length: 1364
 XX Score: 1020.00 Matches: 199
 XX Percent Similarity: 92.86% Conservative: 9
 XX Best Local Similarity: 88.84% Mismatches: 16
 XX Query Match: 87.78% Indels: 0
 XX DB: 24 Gaps: 0
 XX
 XX US-09-245-198a-2 (1-225) x ABR34881 (1-1364)
 XX
 XX 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
 XX 161 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 220
 XX 22 SerTrpAlaThrLeuSerAlaGlnGlnProSerGlnGlnGlnGlnGlnGlnGlnGlnGln 41
 XX 221 AGCGGCGATGCTGCTGCGCCAGAGGCGCTGCCAGAGAGGAGGAGGAGGAGGAGGAGGAG 280
 XX 42 ArgGluProPheGluLeuAsnProGlnThrGlnGlnSerGlnAspValValProPheLeu 61
 XX 281 CAGGACCGCTGGAAGTGAATGCCAGACAGAAAGCAGATCTGCGCTTCTCTGCTG 340

QY 62 GluGlnLeuValArgProArgSerAlaProLysGlyArgGlyAlaArgProArgArg 81
 DB 341 AACGCACTAGTTCGGCTCCGAGAGTCCACCTAAAGCCGGAACACGGGCTCGAAGA 400
 QY 82 AlaIleAlaAlaHisTyrGlnValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
 DB 401 GCGATCGCAGCCCATTTATGAGTTCATCCACGACTGACAGAGGAGGAGGAGGAGGAGGAGT 460
 QY 102 ValAspGlyThrValSerGlyTyrGlnGluThrLysIleAsnSerSerSerProLeuArg 121
 DB 461 GTGGACGGGACAGTGAAGTGGCTGGAGGAAACCAAAACACACCTCCAGCCCTCTGGCC 520
 QY 122 TyrAspArgGlnIleGlyLysPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
 DB 521 TACAAACGCCCAATCGGGAGATTATATAGTACACCGGGCTGGGCTTACTACTGATGCT 580
 QY 142 GlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
 DB 581 CAGTGCACTTTGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 640
 QY 162 ValLeuAlaLeuAlaArgCysLeuGlnGluPheSerAlaThrAlaIleAspSerProGlyPro 181
 DB 641 GTGCTGGCCCTGCGCTGCTGAGAAATCTCAGCCACTGCGGCGAGTTCCTCGGCGCC 700
 QY 182 GlnLeuAlaGlyLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
 DB 701 CAGCTCCGCTCTGCTGCAAGTGTCTGGCTGTGGCTGTGGCCCTGCGGCTCTCTCTGCGG 760
 QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
 DB 761 ATCCGACCCCTCCCTGGGCGCCATCTCAAGGCTGGCCCTTCACCTACTGCGACTG 820
 QY 222 PheGlnValHis 225
 DB 821 TTCAGGTTTCAC 832
 XX
 XX RESULT 8
 XX ID AAV18600 standard; cDNA; 1373 BP.
 XX AC AAV18600;
 XX 21-JUL-1998 (first entry)
 XX
 XX Homo sapiens tumour necrosis factor related ligand (TRELL) gene.
 XX
 XX TRELL; tumour necrosis factor related ligand; tnfr; treatment;
 KW cancer; autoimmune disease; immune system; stimulation; suppression;
 KW graft rejection; ds.
 XX
 XX OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 XX CDS 1..852 /*tag= a
 XX FT /*note= "tumour necrosis factor related ligand"
 XX FT
 XX PN MO9805783-A1.
 XX
 XX 12-FEB-1998.
 XX
 XX 07-AUG-1997; 97WO-US13945.
 XX
 XX 18-MAR-1997; 97US-0040820.
 XX 07-AUG-1996; 96US-0023541.
 XX 18-OCT-1996; 96US-0028515.
 XX
 XX (BIOJ) BIOGEN INC.
 XX (UTGE-) UNIV GENEVA FACULTY MEDICINE.
 XX Browning JL, Chicheportliche Y;
 XX WPI; 1998-145619/13.

DR P-PSDB; AAW47525.
 XX Tumour necrosis factor related ligand - useful for, e.g. treating
 PT cancer, auto-immune disease and immune responses to tissue grafts
 XX
 PS Claim 2: Pages 48-50: 69pp; English.
 XX
 CC The sequence is that encoding human tumour necrosis factor related
 CC ligand (TRELL). TRELL or active fragments can be included with a
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune
 CC diseases or immune responses to tissue grafts, or to stimulate or
 CC suppress the immune system. It is useful to screen for TRELL
 CC receptors, by labelling with a detectable label and screening
 CC compositions for binding. Agents interfering with TRELL-receptor
 CC binding can also be screened for, can then be administered,
 CC optionally with interferon- γ , to induce cell death or
 CC treat, suppress or alter immune responses (especially involving human
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its
 CC receptor. The DNA sequence can be used in gene therapy for
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
 CC autoimmune and inflammatory diseases or inherited genetic disorders,
 CC by introducing into cells, and expressing, therapeutically effective
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
 CC It may also be of use in the preparation of prepare probes for
 CC screening natural/synthetic DNAs for TRELL-encoding sequences
 CC and for antisense therapy.
 CC
 XX
 SQ Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;
 Alignment Scores:
 Pred. No.: 9,12e-84 Length: 1373
 Score: 1020.00 Matches: 199
 Percent Similarity: 92.86% Conservative: 9
 Best Local Similarity: 88.84% Mismatches: 16
 Query Match: 87.78% Indels: 0
 DB: 19 Gaps: 0
 US-09-245-198a-2 (1-225) x AAV18600 (1-1373)
 QY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
 DB 181 CTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 240
 QY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41
 DB 241 AGCGGGGCAATCGCTGCTCCGCCAGAGCCCTGCCAGAGCAGCTGGCGAGAGAGAC 300
 QY 42 ArgGluProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeu 61
 DB 301 CAGGACCCGTCGCACTGAATCCCAACAGAGAAAGCCAGATCCTGGCCCTTCTCTG 360
 QY 62 GluGlnLeuValArgProArgSerAlaProGlyGlyArgLysAlaArgProArgArg 81
 DB 361 AACCGACTAATTCGGGCTCCAGAGTGCACCTAAAGGGCGGAAACACGGGCTCCAGAGA 420
 QY 82 AlaIleAlaIleAlaIleSerGlyValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
 DB 421 GCATATGCGACCCCATTAATGAACTTCATCCAGACTGGAGAGACGAGCGAGCGAGGT 480
 QY 102 ValAspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerProLeuArg 121
 DB 481 GTGGACGGGCACTGACTGGCTGGAGAGAGCCAGAAATCAACAGCTCCACCCCTCTGGCG 540
 QY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuGlyTrpLeuTyrCys 141
 DB 541 TACAACCGCCAGATCGGGAGGTTATAGTACACCGGGCTGGGCTACTACTACTACTACT 600
 QY 142 GluValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
 DB 601 CAGGTCACTTTGATGAGGGAAGGCTGTACTGAACTGAACTGAGTGGTGGATGGT 660
 QY 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSerSerProGlyPro 181

DB 661 GTGCTGGCCCTGGCTGGCTGGAGGAATTCAGCCACTGGCGGCAAGTCCCTGGGCCC 720
 QY 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
 DB 721 CAGCTCCGCTCTGCGCAGGTGTGTGGCTGTGGCCCTGGCGGCAAGGTCTCCCTGGCG 780
 QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
 DB 781 ATCGGACCCCTCCCTGGGCGCATCTCAAGGCTGCCCTCTCTCACTTGGGACTC 840
 QY 222 PheGlnValHis 225
 DB 841 TTCAGAGTTCAC 852
 RESULT 9
 AAX56000
 ID AAX56000 standard; DNA: 1421 BP.
 XX
 AC AAX56000;
 XX
 DT 15-JUL-1999 (first entry)
 XX
 DE Human tumour necrosis factor Apo-3 ligand polynucleotide sequence.
 XX
 KW Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
 KW NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
 KW cancer; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 92..841
 FT /tag= a
 FT /product= "Apo-3 ligand"
 XX
 PN WO919490-A1.
 XX
 PD 22-APR-1999.
 XX
 PF 09-OCT-1998; 98WO-US21407.
 XX
 PR 17-DEC-1997; 97US-0069862.
 PR 10-OCT-1997; 97US-0062037.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Ashkenazi AJ, Marsters SA, Pittl R;
 XX
 DR WPI: 1999-287982/24.
 DR P-PSDB: AAT09369.
 XX
 PT New human Apo-3 ligand (a tumour necrosis factor) homologue
 XX
 PS Claim 18: Fig 1; 74pp: English.
 XX
 CC The present sequence encodes a human tumour necrosis factor (TNF) and
 CC lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
 CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in
 CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and
 CC to induce JNK/SAPK-dependent responses in mammalian cells.
 XX
 SQ Sequence 1421 BP; 281 A; 464 C; 404 G; 272 T; 0 other;
 Alignment Scores:
 Pred. No.: 9,53e-84 Length: 1421
 Score: 1020.00 Matches: 199
 Percent Similarity: 92.86% Conservative: 9
 Best Local Similarity: 88.84% Mismatches: 16
 Query Match: 87.78% Indels: 0
 DB: 20 Gaps: 0
 US-09-245-198a-2 (1-225) x AAX56000 (1-1421)


```

Oy 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGly 21
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 167 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 226
Oy 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluGluLeuThrAlaGluAspArg 41
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 227 AGCGGCGCATGCGTCTCCGCCAGAGCGCTGCCAGAGAGAGCTGGTGGAGAGAGAGC 286
Oy 42 ArgGluProProGlnLeuAsnProGlnThrGlnGluSerGlnAspValValProPheLeu 61
    ::::::::::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 287 CAGGACCGCTGGAACTAAATCCAGACAGAAAGACAGAGCTCGCGCTTTCTG 346
Oy 62 GlnGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgArg 81
    ::::::::::| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 347 AACGACACAGTTTGGCGCTCGCAGAGAGTGCACCTAAAGGCCGAAACACGGGCTGAGGA 406
Oy 82 AlaIleAlaIleAlaIleSerGluValHisProArgProGlyGlnAspGlyAlaGlnAclly 101
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 407 GCGATCGCAGCCCATTAAGATTTCATCCAGACCTGGACAGAGAGCGCAGCAGAGT 466
Oy 102 ValAspGlyThrValSerGlyThrGlnGluThrLysIleAsnSerSerSerProLeuArg 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 467 GTGACGGGACAGTAGTGCTGGGAGAGAGCCAGAAATCAACAGCTCCAGCCCTTGCGC 526
Oy 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 527 TACAACCGCCAGATCGGGGAGATTATAGTCAACCGGGGCTGGGCTCTACTACTGACTG 586
Oy 142 GlnValHisPheAspGlnGlyLysAlaValTyrLeuLysLeuAspLeuLeuValAsnGly 161
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 587 CAGGCGCACTTGTATGAGAGGGGAGAGCTGTACTGAAAGCTGACTGCTGTGATGATG 646
Oy 162 ValLeuAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaIleSerSerProGlyPro 181
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 647 GTGCGCGCTGGCGCTGGAGAAATCTCAGCCACTGGCGGCGAGTTCTCTGGGCC 706
Oy 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 707 CAGCTCGCGCTGTGCGAGGTGCTGTGGCTGTGGCCCTGGCGGAGGCTCCCTCGCGG 766
Oy 202 IleArgThrLeuProTrpAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 767 ATCCGACACCTCCCTCGGCGCATCTCAAGGCTGCCCTTCCTCATTCTCGAGCTC 826
Oy 222 PheGlnValHis 225
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 827 TTCGAGGTTTCAC 838

RESULT 10
AAS03964
ID AAS03964 standard: DNA: 898 BP.
XX
AC AAS03964:
XX
DT 26-SEP-2001 (first entry)
XX
XX Expression vector pDC409-IZ-TWEAK fusion protein-encoding DNA.
XX
KW TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;
KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism; ds;
KW corneal graft neovascularisation; psoriasis; metastatic condition;
KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
KW preneoplastic condition; myocardial angiogenesis; wound granulation;
KW scleroderma; vascular adhesion; telangiectasia; ischemia; human;
KW atherosclerotic plaque neovascularisation; coronary atherosclerosis;
KW peripheral atherosclerosis; pDC409-IZ-TWEAK; TWEAK receptor; TWEAKR;
KW fusion protein.
XX
OS Homo sapiens.
OS Synthetic.
XX

```

```

FH Key Location/Qualifiers
FT CDS 52..873
FT /tag= a
FT /product= "Fusion protein comprising a growth hormone
FT leader, a leucine zipper multimerisation
FT domain, and human TWEAK extracellular
FT domain"
XX
XX MO200145730-A2.
XX
XX 28-JUN-2001.
XX
XX 19-DEC-2000; 2000WO-US34755.
XX
XX 20-DEC-1999; 99US-0172878.
XX
XX 10-MAY-2000; 2000US-0203347.
XX
XX (IMMV ) IMMUNEX CORP.
XX
XX WILEY SR;
XX
XX WPI: 2001-417975/44.
XX
XX P-PSDB: AAU03499.
XX
XX
XX Modulating angiogenesis in a mammal for treating diseases mediated by
XX angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
XX peripheral tissue, by administering antagonist or agonist of TWEAK
XX receptor
XX
XX Example 1: Page 39-40; 46pp; English.
XX
XX The sequence represents a DNA from the expression vector
XX pDC409-IZ-TWEAK, which encodes a fusion protein comprising a growth
XX hormone leader, a leucine zipper multimerisation domain, and the
XX extracellular domain of human TWEAK. The fusion protein was used in
XX the isolation of human TWEAK receptor (TWEAKR)-expressing clones
XX from a COS cell human cDNA library. The TWEAKR protein is a
XX member of the tumour necrosis factor (TNF) family and induces
XX angiogenesis. TWEAKR may therefore be used to screen for and develop
XX TWEAKR agonists and antagonists for the modulation of angiogenesis, to be
XX used in the treatment and diagnosis of human disease. The disorders
XX mediated by angiogenesis include ocular disorders characterised by ocular
XX neovascularisation such as diabetic retinopathy, neovascular glaucoma,
XX retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
XX rubeosis, uveitis, macular degeneration and corneal graft
XX neovascularisation, and inflammatory diseases such as arthritis,
XX rheumatism and psoriasis. Other treatable diseases include malignant and
XX metastatic conditions such as sarcomas and carcinomas, benign tumours and
XX preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
XX scleroderma, vascular adhesions, atherosclerotic plaque
XX neovascularisation, telangiectasia, wound granulation, coronary
XX atherosclerosis, peripheral atherosclerosis and ischemia.
XX
XX Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other:
XX
XX
XX Alignment Scores:
XX Pred. No.: 1.06e-77 Length: 898
XX Score: 951.00 Matches: 184
XX Percent Similarity: 93.24% Conservative: 9
XX Best Local Similarity: 88.89% Mismatches: 14
XX Query Match: 81.84% Indels: 0
XX DB: 22 Gaps: 0
XX
XX US-09-245-198a-2 (1-225) x AAS03964 (1-898)
Oy 19 SerLeuGlySerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluGluLeuThrAla 38
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 250 AGTTTGGGAGACCGGCGCATCGCTGCCGCCAGAGAGCTGCCAGAGAGAGCTGGTGC 309
Oy 39 GlnAspArgArgGluProProGlnLeuAsnProGlnThrGlnGluSerGlnAspValVal 58
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 310 GAGGAGACAGACCGCTCGGAATCTCCGCCAGACAGAAAGAACCCAGAGATCTCTGCG 369

```

Oy	59	ProheuleuglugluLeuValaLarProAraArgSerSeraProLyGlyAArgysLaArg	78
Db	370	CGTTCTCTGAAACGACAGTTCGGCCCTCGAGAAAGTCACTPAAGGCGGAAAAACACGG	429
Oy	79	ProArgaArgLaIleAlaAlaHisTyrGluValaHisProArgProGlyGlnAspGlyAla	98
Db	430	GCTCGAAGAGAGATCGACGCCCATTTATGAAGTTCATCCACGACCTGACAGAGGACGGACG	489
Oy	99	GlnAlaGlyValaAspGlyThrValaSerGlyTyrPgluGluThrLysIleAsnSerSerSer	118
Db	490	CAGGACAGTGTGGACGGGACAGTAgAGTGGCTGGAGGAAGCCAGAAATCAACAGCTCCAGC	549
Oy	119	ProLeuArgTyrAsaArgGlnIleGlyGluPheThrValaIleArgAlaGlyLeuTyrTyr	138
Db	550	CCCTGGCGCATCAACACCCGATCGGGAGATTATATGTCACCCGGGCTGGCTACTACAC	609
Oy	139	LeuTyrCysGlnValaHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuLeu	158
Db	610	CTGACCTGTACAGGAGCTTTGATGAGGGGAAGGCTGTACTCTGAAGCTGACATTGGTG	669
Oy	159	ValasnGlyValaLeuAlaLeuAlaArgCysLeuGluGlnPheSerAlaThrAlaIleSerSer	178
Db	670	GTGATGTGTGTGGCGCCCTGCGCTGCTGAGGAATTTCTCAGCCACTGGGCCAGTTCC	729
Oy	179	ProGlyProGlnLeuArgLeuCysGlnValaSerGlyLeuLeuProLeuArgProGlySer	198
Db	730	CTCGGGCCCCAGCTCCGCTCTCCAGAGTGTCTGCGCTTGGCCCTGGCGCCAGGGTCC	789
Oy	199	SerLeuArgTyrIleAArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyr	218
Db	790	TCCCTGGGAGATCGCACCTCCCTCGGAGCCATCTCAAGAGTGGCCCTTCTCTCACTAC	849
Oy	219	PheGlyLeuPheGlnValaHis	225
Db	850	TTTCGACCTCTTCACAGTTTAC	870
RESULT 11			
AAAX23424			
ID	AAAX23424	standard; DNA: 1030 BP.	
AC	AAAX23424:		
XX	18-JUN-1999	(first entry)	
DT			
XX	Human TNRL3 DNA.		
DE			
XX	Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;		
KW	developmental abnormality; gestational abnormality; prostate cancer;		
KW	AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease		
KW	cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;		
KW	apoptosis; human; ss.		
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	1..627	
FT		/tag= a	
FT		/product= "TNRL3"	
XX			
PN	MO9911791-A2.		
XX			
PD	11-MAR-1999.		
XX			
PE	04-SEP-1998;	98MO-US18393.	
XX			
PR	05-SEP-1997;	97US-0924634.	
XX			
PA	(UNIM) UNIV WASHINGTON.		
XX			
PI	Chauchary PM.		
XX			
XX	WPI: 1999-205191/17.		
OR	P-PSDB: AAM93590.		
OR			

XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PI developmental or gestational abnormalities
XX
PS Example VII; Fig 13A; 156pp; English.

This invention describes isolated Tumor Necrosis Factor (TNF) family receptor polypeptides: AP04, AP06, AP08 and AP09 or their active fragments, and isolated TNF related ligands 1 and 3 (TNFR1 and TNFR3) or their active fragments. AP04 is useful for diagnosing prostate cancer by determining levels of AP04 in an individual. Prostate cancer can also be treated using AP04 selective binding agents linked to a therapeutic moiety. AP04 polypeptides are also useful for identifying selective binding agents, useful in diagnosis/treatment of disease by binding of agents to the polypeptide/active fragment which is extracellular, or expressed on the cell surface. The binding is preferably performed *in vivo*. AP04 polypeptides/active fragments are also useful for screening for agonists and antagonists by binding and observing the change in AP04 activity. Effective pharmacological agents useful in diagnosis or treatment of disease are also identified using AP04 polypeptides/active fragments and AP04 signal transducer molecules that specifically interact with a cytoplasmic domain of AP04 and detecting a change in level of AP04 activity. The method is performed *in vivo* or *in vitro*. AP04 polypeptides are all useful as immunogens for preparing antibodies. AP04 is also useful for diagnosis/treatment of developmental or gestational abnormalities. AP08 was transfected to human breast carcinoma cell line MCF-7, and induced apoptosis.

Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;

Alignment Scores:	
Prod. NO.:	1,26e-77
Score:	951.00
Percent Similarity:	92.79%
Best Local Similarity:	88.94%
Query Match:	81.84%
DB:	20
Length:	103
Matches:	185
Conservative:	8
Mismatches:	15
Indels:	0
Gaps:	0

US-09-245-198A-2 (1-225) x AAX23424 (1-1030)

QY 18 ValSerLeuGlySerTTPalaIthLeuSerAlaIactInIuProSerGInIuGluLeuThr 37
Db 1 GTCACATTTTGGGAGCCGGGATTCGCTGTCCGCCAGGAGCCGTCCCGGAGGAGCTGGTG 60
QY 38 AlaGluAspArgArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspVal 57
Db 61 GCAGAGGAGAGACAGACGACCCGTGGAACTGAATCCCAAGACAGAAAGACCAAGATCCT 120
QY 58 ValProPheLeuGluGlnLeuValIatArgProArgSerAlaPolySgiIatArglyAla 77
Db 121 GCGCCTTTTCGAAACCGCACTAGTTCCGGCTCGCGAAGATGCACCTAAAGCCGGGAAAACA 180
QY 78 ArgProArgArgAlaIleAlaAlaHisIstArgIuValHisProArgProGluGlnAspGly 97
Db 181 CGGGCTCCAAAGAGCATGCGACGCCATTATGAAGTTTATCCACAGCACTGGACAGACGAGA 240
QY 98 AlaGlnAlaGlyValAspGlyThrValSerGlyTTPGluGluThrIysIleAsnSerSer 117
Db 241 GCGCAGCGCAGGTGTGGACGGACAGTGAATGGCTGGGAGGAGCAAGCAATCAACAGCTCC 300
QY 118 SerProLeuAlaGlyTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyr 137
Db 301 AGCCCTTCGCCCTACTACCGGCCAAGTGGGGAGATTTPATGTCACCCCGGGGTGGACTTCAC 360
QY 138 TyrLeuTyrCysGlnValHisPheAspGluGlyIysAlaValTyrLeuLysIleAspLeu 157
Db 361 TACCTGTACTGTCAAGTGCACATTGTGATGAGGGAGAACGCTGTACCTGAAGCTGGACTTG 420
QY 158 LeuValAsnGlyValLeuAlaLeuAlaArgCysLeuGluGluPheSerAlaThrAlaIAser 177
Db 421 CTGGAGATAGGTGTGTGTCGTCCTCGCTGCTGGAGGAAATTCACAGCACTGGCGGCAGT 480

PF 23-MAR-2001; 2001WO-US09231.
 XX 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI: 2001-656860/75.
 XX
 DR WPI: 2001-656860/75.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 XX
 PS Claim 1: SEQ ID NO 15892; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
 CC sequences (AB161737-AB162072).
 CC (AB161737-AB162072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 SQ Sequence 978 BP; 289 A; 261 C; 238 G; 190 T; 0 other;
 Alignment Scores:
 Pred. No.: 0.663 Length: 978
 Score: 110.50 Matches: 49
 Percent Similarity: 41.62% Conservative: 33
 Best Local Similarity: 24.87% Mismatches: 81
 Query Match: 9.51% Indels: 34
 DB: 23 Gaps: 9
 US-09-245-198a-2 (1-225) x ABL21473 (1-978)
 OY 41 Argatggluprprogluleuansbproglunhrngluserglnaspyalvalpro-ph 60
 DB 450 CAAAGAGAGCCCTGCACCACTTCACCA-----CCGTCCAGAAATGATTCCTCCGCA 500
 OY 60 eleuclnglnleuvalargproargarqser-----AaProlysglyArglyAlaLar 78
 DB 501 TCGCCACCTCTAGTCCGCGAAAGTGAAATCTCTTCTTCAGCCAGATCCGAGGACTCGAG 560
 OY 78 gproargatgalaialaialaiahistyrclnvalahisproargproglnglnaspyal 98
 DB 561 GCCA-----GCAGCCCATTTCCACTTGCAGCAGCAGCGCGCGCACCAAGGAG 608
 OY 98 aglnaiaaglyalaspqlythvalsercltyrprnglunhrhlyleasnserserse 118
 DB 609 TATG---GGCTACCATGCGATATGATACATAGAAATGATACAGAGAAACTCTTATCA 665
 OY 118 rProleuargtyrargparglnilleglyclunphthrvalilearglaaglyleuertyty 138
 DB 666 GGGA---CACTTTCAAACGCGGATGCGCTCTTTCAGCGGACCAATACAGGCTATATTA 722
 OY 138 rleuTyrcysglnvalahisphaspqluglylsalvaltyrleuLysleuaspLeule 158
 DB 723 CGTATACGCCCGCATATGCTACACAACTCCGACGAC-----759
 OY 158 uValasnglyalLeuala-----leuargyLysleuGlunph 171
 DB 760 -CAGAACCGATTATTCGTTCTTCAAGAGACACATTCCTCGCAGTGGTTGAAC----- 813
 OY 171 eSerlathrAlaAlaSerSerProglyProclnleuargLysGlnvalSerGlyle 191
 DB 814 -----ACGGTGGCCACCAACATGCGACATTAAGGTGACACACTTCCACAGAGGTGTCT 866

OY 191 uleuProleuargProglySerleuarglileargThleu-----ProTrpAl 208
 DB 867 GATCCACCTGGAGACGAACGAGAGATCATCTGAAGGACATTCACACGATCGCAATGC 926
 OY 208 ahisLeuLysAlaAlaProphleuThrTyRPhcglyleupheglnval 224
 DB 927 AGTCTCGCGGAGGAGGAACCAACCGAAGCTACTTGTGCACTCTCAAGGTG 975
 RESULT 15
 AA41377
 ID AA41377 standard; cDNA; 1630 BP.
 AC AA41377;
 XX
 DT 08-OCT-1998 (first entry)
 XX
 DE NF-kB receptor activator RANK ligand (RANKL) encoding cDNA.
 XX
 KW RANK; necrosis factor-kappa B; NF-kB; receptor activator; human;
 KW immune response; inflammatory response; toxic shock; sepsis;
 KW RANKL; RANK ligand; tumour necrosis factor; TNF; mouse; ss.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT CDS 3..887
 FT /tag= a
 FT /product= "murine RANKL (ligand for RANK)"
 XX
 PN WO9828426-A2.
 XX
 PD 02-JUL-1998.
 XX
 PF 22-DEC-1997; 97WO-US23775.
 XX
 PR 14-OCT-1997; 97US-0064671.
 PR 23-DEC-1996; 96US-0059978.
 PR 07-MAR-1997; 97US-0813509.
 XX
 PA (IMMUNEX CORP.
 XX
 PI Anderson DM, Galibert LJ, Maraskovsky E;
 XX
 DR WPI: 1998-377657/32.
 DR P-PSDB: AAM69956.
 XX
 PT New isolated ligand for receptor activator of NF-kappa B - used to
 PT develop products for augmenting an immune response for inhibiting an
 PT inflammatory response and for protection of cells
 XX
 PS Claim 25; Pages 55-57; 80pp; English.
 XX
 CC This cDNA encodes a murine RANKL, a ligand for the RANK (receptor
 CC activator of necrosis factor-kappa B (NF-kB)) polypeptide. RANK is a
 CC member of the tumour necrosis factor (TNF) family. A soluble RANK
 CC may be used for inhibiting activation of NF-kB, by contacting a cell
 CC expressing membrane-associated RANK with a soluble RANK which binds to
 CC RANK ligand (RANKL). RANKL polypeptides can activate RANK and can be
 CC used to induce maturation of dendritic cells and enhance their
 CC allo-stimulatory capacity, thereby augmenting an immune response. The
 CC soluble RANK polypeptide composition may also be used for regulating an
 CC immune or inflammatory response. Inhibition of NF-kB by RANK antagonists
 CC may be useful in ameliorating negative effects of an inflammatory
 CC response that result from triggering of RANK, e.g. in treating toxic
 CC shock or sepsis, graft-versus-host reactions, or acute inflammatory
 CC reactions. They can also be used in adjunct therapy for disease
 CC characterised by neoplastic cells that express RANK. RANKL polypeptides
 CC can also be used to identify inhibitors of RANK and thus inhibitors of
 CC an inflammatory response, and also for protecting RANK-expressing cells
 CC from the negative effects of chemotherapy or the presence of high levels
 CC of TNF-alpha. The products can also be used for detection and drug
 CC screening.
 XX

Sequence 1630 BP; 436 A; 355 C; 379 G; 460 T; 0 other;

Alignment Scores:

Pred. No.:	2.95	Length:	1630
Score:	106.50	Matches:	67
Percent Similarity:	38.04%	Conservative:	38
Best Local Similarity:	24.28%	Mismatches:	106
Query Match:	9.17%	Indels:	65
DB:	19	Gaps:	13

US-09-245-198a-2 (1-225) x AAV41377 (1-1630)

```

QY      3 SerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuValValSerLeuGlySer 22
      |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      78 TCCATGTTCTGCGCCCTCGGGGCTGGGACTGGGCGAGCTGGCCAGGCTGCAGCATCGCTCTG 137
QY      23 TrpAlaThrLeuSerAlaGln---GluProSerGlnGluGluLeuThrAlaGluAspArg 41
      :::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      138 TTCCTGTACTTTCGAGCGGAGATGATCCTTAACAGA-----ATATCAGAAGACAGC 188
QY      42 -----ArgGluProProGluLeuAsnProGlnThr 51
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      189 ACTCACTGCTTTATATGAATCCTGAGACTCCATGAACGAGATTGCGAGACTCGACT 248
QY      52 GluGluSerGlnAspValAlaPro----- 59
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      249 CTGGAGAGTGACAGACACACTACTGCTCTCGAGAGAGATGAACACAGCTTTCAGGGG 308
QY      60 -----PheLeuGlnGlnLeuValAlaArgProArgArg-----SerAlaProLys 73
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      309 GCCGTGCAGAAAGAACTGACACATGTTGGGCCACAGGCTTCTCAGAGAGCTCCAGCT 368
QY      74 -----GlyArgLysAlaArgProArgArgAlaIleAla 84
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      369 ATGATGGAAGGCTCATGTTGGATGTGGCCACGAGCAAGCCTGAGGCCCAAGCATT 428
QY      85 AlaHisTrpGlnValHisProArgProGlyGlnAspGlyValGlnAlaGlyValAspGly 104
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      429 GCACACTCTACATCAATGCTGCCAGCATCCATGGGTTCCCAT-----AAAGTC 479
QY      105 ThrValSerGlyTrpGluGlu-----ThrLysIleAsnSerSerProLeu 120
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      480 ACTCTGTCTCTTGTGTACACGATCGAGGCTGGGCCAAGATCTCTAACATGACG----- 533
QY      121 ArgTyrAspArgGlnIleGlyLupheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      534 -----TTAAGCAACGAAAGAACTAAGGTTAACCAAGATGCTTCTATTACTCTAC 584
QY      141 -----CysGlnValHisPheAspGluGlyLysAlaVal-----TyrLeuLys 154
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      585 GCCAACATTTGGCTTTCGGCATCATGAACATCGGGAAGCGTACCTACAGACTATCTTCAG 644
QY      155 LeuAspLeuLeuVal-----AsnGlyValLeuAlaLeuArgCysLeuGluGlu 170
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      645 CTGATGGTGTATGTCGTTAAACACAGCATCAAAATCCCAAGTCTCATTAACGTGATGAAA 704
QY      171 PheSerAlaThrAlaAlaSerSerProGlyProGlnLeuArgLeu-----CysGlnVal 188
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      705 GGAGGGAGCAGCAAAACTGTCGGGCAATTCGAATTCACCTTTTATCCATAAATGTT 764
QY      189 SerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeuProTrpAla 208
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      765 GGGGATTTTCAAGCTCGAGCTGTGAGAAATTTGCAATTCAGTGTCCAAACCCCTTCC 824
QY      209 HisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 224
      |||::: |||::: |||::: |||::: |||::: |||::: |||:::
DB      825 CTGCTGGATCCGATCAAGATGCGACGTACTTGGGGCTTTCAAAAGTT 872

```

Search completed: March 31, 2003, 02:40:55
 Job time : 241.957 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 00:53:03 ; Search time 2837.48 Seconds

(without alignments)
2307.731 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162
Sequence: 1 VLISGLALACGLLVVSL.....PWAHLKAPFLTYFLQVH 225

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008
Listing first 45 summaries

Command line parameters:
-MODE=frame+;p2n.model -DEV=xlh
-Q=/gen2.1/USPTO/US09245198/runat_24032003_163553_25950/app.query.fasta.1.846
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blsum62 -TRANS=human40.cdi -LIST=45
-LOCALALIGN=200 -THR.SCORE=DCT -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US9245198 -ECGN.1.1.3077 -runat_24032003_163553_25950 -NCPU=6 -ICPU=3
-NO_XLPTX -NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMECUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
1: GenEmbl:*
2: gb_da:*
3: gb_hg:*
4: gb_in:*
5: gb_om:*
6: gb_ov:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_da:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mus:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_hg_hum:*
31: em_hg_inv:*
32: em_hg_other:*
33: em_hg_mus:*
34: em_hg_pln:*
35: em_hg_rod:*
36: em_hg_mam:*
37: em_hg_vrt:*
38: em_sy:*
39: em_hgo_hum:*
40: em_hgo_mus:*
41: em_hgo_other:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1162	100.0	1168	10	AF030100	AF030100 Mus muscu
2	1020	87.8	1236	6	AR140407	AR140407 Sequence
3	1020	87.8	1306	9	AF030099	AF030099 Homo sapi
4	1020	87.8	1353	6	AX201324	AX201324 Sequence
5	1020	87.8	1368	9	AF055872	AF055872 Homo sapi
6	951	81.8	898	6	AX180714	AX180714 Sequence
7	945.5	81.4	1651	9	BC019047	BC019047 Homo sapi
8	602.5	51.9	203083	2	AC069459	AC069459 Mus muscu
9	602.5	51.9	234182	10	AL603707	AL603707 Mouse DNA
10	549.5	47.3	177703	2	AC016876	AC016876 Homo sapi
11	549.5	47.3	215795	2	AC127470	AC127470 Pan trogl
12	545	46.9	17555	2	AC130192	AC130192 Sus scrof
13	544.5	46.9	161428	2	AC126925	AC126925 Canis fam
14	536.5	46.2	18792	2	AC119115	AC119115 Rattus no
15	503	43.3	153553	2	AC126921	AC126921 Bos tauru
16	425.5	36.6	184026	2	AC098923	AC098923 Rattus no
17	207	17.8	203281	2	AC126237	AC126237 Canis fam
18	113.5	9.8	1656	3	AB073865	AB073865 Drosophi1
19	110.5	9.5	1221	3	AY119233	AY119233 Drosophi1
20	110.5	9.5	2159	3	AF521176	AF521176 Drosophi1
21	109.5	9.4	892	5	GGA243435	AJ243435 Gallus ga
22	109.5	9.4	108967	9	AL353138	AL353138 Human DNA
23	108.5	9.3	42210	1	SC1C2	AL031124 Streptomy
24	107.5	9.3	13243	1	AE004602	AE004602 Pseudomon
25	107	9.2	178262	2	AC061974	AC061974 Homo sapi
26	107	9.2	179383	2	AC060789	AC060789 Homo sapi
27	106.5	9.2	1630	6	AR156433	AR156433 Sequence
28	106.5	9.2	1630	6	AR164147	AR164147 Sequence
29	106.5	9.2	1630	6	AX147987	AX147987 Sequence
30	106.5	9.2	1694	6	AX451897	AX451897 Sequence
31	106.5	9.2	2225	10	AF019048	AF019048 Mus muscu
32	106.5	9.2	4034	6	AX200995	AX200995 Sequence
33	106.5	9.2	4034	6	AX267730	AX267730 Sequence
34	106.5	9.2	4894	6	AX200993	AX200993 Sequence
35	106.5	9.2	4894	6	AX267728	AX267728 Sequence
36	106.5	9.2	6976	6	AX200996	AX200996 Sequence
37	106.5	9.2	157988	6	AX267731	AX267731 Sequence
38	106.5	9.1	177485	2	AC096533	AC096533 Homo sapi
39	106	9.1	207585	2	AC090975	AC090975 Homo sapi
40	106	9.1	222037	2	AC073379	AC073379 Mus muscu
41	106	9.1	222037	2	AC073801	AC073801 Mus muscu
42	105.5	9.1	4412	9	AY070219	AY070219 Homo sapi
43	105.5	9.1	6305	9	HMU1YTOXBB	LI1016 Homo sapien
44	105.5	9.1	24526	6	AX472617	AX472617 Sequence
45	105.5	9.1	62370	2	AL451008	AL451008 Homo sapi

RESULT 1

ALIGNMENTS

AF030100
LOCUS AF030100 1168 bp mRNA linear ROD 20-DEC-1997
DEFINITION Mus musculus TWEAK mRNA, partial cds.
ACCESSION AF030100
VERSION AF030100.1 GI:2707220
KEYWORDS
SOURCE Mus musculus.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 1168)
AUTHORS Chicheportiche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H., Hession,C., Garcia,I. and Browning,J.L.
TITLE TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis
JOURNAL J. Biol. Chem. 272 (51), 32401-32410 (1997)
MEDLINE 98070415
PUBMED 9405449
REFERENCE 2 (bases 1 to 1168)
AUTHORS Chicheportiche,Y., Blixler,S., Flizard,R. and Browning,J.
TITLE Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA
JOURNAL
FEATURES
source location/Qualifiers
1. .1168
/organism="Mus musculus"
/db_xref="taxon:10090"
/cell_type="peritoneal macrophages"
/note="ligand in the TNF family; secreted protein"
/codon_start=2
/product="TWEAK"
/protein_id="AAC53517.1"
/db_xref="GI:2707221"
/translation="VLSTGLAPRALHSLAALVLSGSMATLSAQSSELTAEAREP
PELNPOTBESQDVPFLBOLVPRRSAPKRRKRRPRRAIAHIEVPRQDQAGY
DGVNSGMEETKINSSSPRLRDQIGETVIRALVLYLCVHDEKAVYLKIDLLVN
GLVALRCLEFEFATAASSPQPLRLCOVSGSLPLRGSSSLRITPLWAMHKAPFLTY
FGLEOVH"
BASE COUNT 242 a 360 c 298 g 268 t
ORIGIN
Alignment Scores:
Pred. No.: 1,47e-79 Length: 1168
Score: 1162.00 Matches: 225
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 10 Gaps: 0
US-09-245-198a-2 (1-225) x AF030100 (1-1168)
Qy 1 ValLeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValSerLeu 20
Db 2 GTGCTGAGCCTGGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 61
Qy 21 GlySerTrpAlaThrLeuSerAlaGlnGlnProSerGlnGlnGlnGlnGlnGlnGln 40
Db 62 GGGAGTGGGCAAGCTGTGTGCCAGAGAGCTTCTCAGAGAGAGCTGACAGAGAGAC 121
Qy 41 ArgArgGluProProGluLeuAsnProGlnThrGlnGlnSerGlnAspValValProPhe 60
Db 122 CGCGGGAGAGCCCTCGAATCGAATCCCGACAGAGAAAGCCAGATGTGGTACCTTTC 181
Qy 61 LeuGlnGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArg 80
Db 182 TTGGAAACACTAGTCCGGCTCGAAGAGTGCTCTAAAGCCGGAAGCGCGGCTTCGC 241
Qy 81 ArgAlaIleAlaIleAlaHsITyrgLValHsIProArgProGlyGlnAspGlyAlaGlnAla 100
Db 242 CGAGCTATTGGACGCCCTTATGAGCTTCACTCTGCCAGAGAGAGATGGAGCACAAGCA 301
Qy 101 GlyValAspGlyThrValSerGlyTTPGluGluThrLysIleAsnSerSerProLeu 120

|||||
Db 302 GGTGTGATGGAGACAGTACGTGGCTGGAGAGACCAAAATCAACAGCTCCAGCCCTCTG 361
Qy 121 ArgTyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyr 140
Db 362 CGTACGACCGCCAGATGGGGAAATTTACAGTCATCAGGGCTGGCTCTACCTGTAC 421
Qy 141 CysGlnValHsIspheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsn 160
Db 422 TGTCAAGTGCACCTTGTATGAGGAAAGCGTCTACTGAAAGCTGGAGCTGTGGTAAAC 481
Qy 161 GlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerProGly 180
Db 482 GGTGTGCTGGCCCTGCGCTGCTGAGAAATTCACACACAGACGACACCTCTCCGCG 541
Qy 181 ProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeu 200
Db 542 CCCAGCTCCGTTGTGTGCCAGGTGTCTGGCTGTGGCTGGGCGCAGGGCTTCCCTT 601
Qy 201 ArgIleArgThrLeuProTrpAlaHsIleLeuLysAlaAlaProPheLeuThrTyrPheGly 220
Db 602 CGGATCGGCAACCTCCCTGGGCTCATCTTAAGGCTGCCCTTCTTACCTACTTGA 661
Qy 221 LeuPheGlnValHs 225
Db 662 CTCTTCAAGTTCAC 676
RESULT 2
ARI140407
LOCUS ARI140407 1236 bp DNA linear PAT 16-JUN-2001
DEFINITION Sequence 1 from patent US 6207642.
ACCESSION ARI140407
VERSION ARI140407.1 GI:14482903
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Willey,S.R.
TITLE Member of the TNF family useful for treatment and diagnosis of disease
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;
FEATURES
source location/Qualifiers
1. .1236
/organism="unknown"
BASE COUNT 225 a 416 c 358 g 237 t
ORIGIN
Alignment Scores:
Pred. No.: 9.38e-69 Length: 1236
Score: 1020.00 Matches: 199
Percent Similarity: 92.86% Conservative: 9
Best Local Similarity: 88.84% Mismatches: 16
Query Match: 87.78% Indels: 0
DB: 6 Gaps: 0
US-09-245-198a-2 (1-225) x ARI140407 (1-1236)
Qy 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValSerLeuGly 21
Db 76 CTGGGCTGGGCGCTGGGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 125
Qy 22 SerTrpAlaThrLeuSerAlaGlnGlnProSerGlnGlnGlnGlnGlnGlnGln 41
Db 136 AGCGGGCATCGCTGTCCGCCAGAGCTTGGCCAGAGAGAGCTGTGGCAGAGAGAGAC 195
Qy 42 ArgGluProProGluLeuAsnProGlnThrGlnGlnSerGlnAspValValProPheLeu 61
Db 196 CAGAGCCCGTCGGAATCGAATCCCGACAGAGAAAGCAAGATCCTGGCCTTCTCTG 225
Qy 62 GluGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgArg 81
Db 256 AACGACTAGTTCGCTCGAAGAGTGCACTTAAGGCGGAAACACGCGGCTCGAAGA 315

QY	82	AlaIleAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly	101
Db	316	GGGATTCGAGCCCATTAAGAAGTTATCATCCAGACCTGGACAGACGAGCGAGGAGGT	375
QY	102	ValAspGlyThrValSerGlyTyrPgluGluThrTrpHisIleAsnSerSerProLeuArg	121
Db	376	GGGACGGGACAGTGAAGTGGCGTGGAGGAGGACCAATCAACAGCTCCAGCCCTTGGCC	435
QY	122	TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys	141
Db	436	TACAAACCCCGAGATCGGGAGAGTTTATGTACACCCGGCGTGGCTTACTACTGTACTGT	495
QY	142	GlnValHisPheAspGluGlyLysAlaValValTyrLeuLysIleAspLeuLeuValAsnGly	161
Db	496	CAGGGCGACCTTGAATGAAGGGAGAGGCTGTCTTACCTGGAAGCTGGACTTCTGGTGAATGT	555
QY	162	ValLeuAlaLeuArgCysGluLeuGluIuPheSerAlaThrAlaAlaSerSerProGlyPro	181
Db	556	GGCGTGGCCCGTGGCTGGCTGGAGGAATTCACGCACTCGCGGAGTTCCCTCGGGGCC	615
QY	182	GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg	201
Db	616	CAGCTCCCGCTCTGCGAGGTGTCTGGCGTGTGGCCCTGGCGGCGAGGGTCTCCCTCGCGG	675
QY	202	IleArgThrIleProThrAlaHisIleuLysAlaAlaProPheLeuThrTyrPheGlyLeu	221
Db	676	ATCCGACACCTCTCCCTGGGCCCCATTCGAAGCTGCCCTTCCTTCCACTTGGACTC	735
QY	222	PheGlnValHis 225	
Db	736	TTCCAGGTTTCAC 747	
RESULT 3			
AF030099		1306 bp mRNA	linear PRI 20-DEC-1997
LOCUS	AF030099	Homo sapiens TWEAK mRNA, complete cds.	
DEFINITION	AF030099		
ACCESSION	AF030099.1	GI:2707218	
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
MEDLINE			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
CDS			

[illegible]

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1353)
Ashkenazi, A.J., Goddard, A., Godwin, P.J., Gurney, A.L., Hillan, K.J., Marsters, S.A., Pap, J., Pitt, R.M., Roy, M.A., Smith, V., Stone, D.M., Watanabe, C.K. and Wood, W.I.
Compositions and methods for the treatment of tumour
Patent: WO 0153486-A 3 26-Jul-2001.

TITLE
JOURNAL
Genentech, Inc. (US)
Location/Qualifiers

FEATURES
source 1..1353
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 257 a 443 c 389 g 264 t
ORIGIN

Alignment Scores:
Pred. No.: 1 03e-68 Length: 1353
Score: 1020.00 Matches: 199
Percent Similarity: 92.86% Conservative: 9
Best Local Similarity: 88.84% Mismatches: 16
Query Match: 87.78% Indels: 0
DB: 6 Gaps: 0

US-09-245-198a-2 (1-225) x AX201324 (1-1353)

QY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValSerLeuGly 21
Db 133 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 192
QY 22 SerTTPAlaThrLeuSerAlaGlnGluProSerGlnGluGluLeuThrAlaGluAspArg 41
Db 193 AGCGGGGACATCGCTGCCGAG 252
QY 42 ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeu 61
Db 253 CAGGACCCGTCGAGACTGATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312
QY 62 GluGlnLeuValArgProArgSerAlaProLysGlyArgLysAlaArgProArgArg 81
Db 313 AACCGACTATGCGGCTCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 372
QY 82 AlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAla 101
Db 373 GCGATCGACCCATATGAGTTATCATCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432
QY 102 ValAspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerSerProLeuArg 121
Db 433 GTGGACGGGACATGAGTGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 492
QY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
Db 493 TACAAACCGCAGATCGGGAG 552
QY 142 GluValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
Db 553 CAGGTGACATTTATGATGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 612
QY 162 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleAsnSerProGlyPro 181
Db 613 GTGCTGGCGCTGGGCGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 672
QY 182 GlnLeuArgLeuGlyGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
Db 673 CACCTCGCGCTGGCAG 732
QY 202 IleArgThrLeuProTrpAlaHisLeuLysAlaIleAlaProPheLeuThrTyrPheGlyLeu 221
Db 733 ATCCGACACCTCCCTGGGCGCATCTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 792
QY 222 PheGlnValHis 225
Db 793 TTCACAGGTTCAC 804

RESULT 5
AF055872 1368 bp mRNA linear PRI 04-MAY-1998
LOCUS
AF055872 Homo sapiens Ap03/DR3 1ligand (APO3L) mRNA, complete cds.

DEFINITION
AF055872.1 GI:3108230
ACCESSION
AF055872.1
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1368)
Marsters, S.A., Sheridan, J.P., Pitt, R.M., Brush, J., Goddard, A. and Ashkenazi, A.
Identification of a ligand for the death-domain-containing receptor

TITLE
JOURNAL
Curr. Biol. 8 (9), 525-528 (1998)
MEDLINE
9828335
PUBMED
9560343
REFERENCE
2 (bases 1 to 1368)
Marsters, S.A., Sheridan, J.P., Pitt, R.M., Brush, J., Goddard, A. and Ashkenazi, A.
Direct Submission
Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way, South San Francisco, CA 94080, USA
Location/Qualifiers

FEATURES
source 1..1368
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/map="17p13"
1..1368
/gene="APO3L"
58..807
/gene="APO3L"
/function="binds to the death-domain-containing receptor
Ap03/DR3"
/note="TNF homolog"
/codon_start=1
/product="Ap03/DR3 ligand"
/protein_id="AAC39724.1"
/db_xref="GI:3108231"
/translation="MAKRSQRRCRGEPEGALLPLALGGLACGLILLAVSL
GSRASLSAQEPAOEELVABEDDPSLENPQTEESODPAPFLNLRVPRRSAPGRKTR
ARRAIAHAHYEVRPEQDGAQGVDTGVSQGEARINSSPLRYNIOIEFTVTRAGL
YLYICOVHDEKAVYLRKLDLVGVALRCLSEESATRAASLSGLQLRLQVSGLLAL
RGGSSRLRTLRMAHLKAPFLDYGLPQVH"

gene
CDS

BASE COUNT 272 a 443 c 389 g 264 t
ORIGIN

Alignment Scores:
Pred. No.: 1 05e-68 Length: 1368
Score: 1020.00 Matches: 199
Percent Similarity: 92.86% Conservative: 9
Best Local Similarity: 88.84% Mismatches: 16
Query Match: 87.78% Indels: 0
DB: 9 Gaps: 0

US-09-245-198a-2 (1-225) x AF055872 (1-1368)

QY 2 LeuSerLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuValValSerLeuGly 21
Db 133 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 192
QY 22 SerTTPAlaThrLeuSerAlaGlnGluProSerGlnGluGluLeuThrAlaGluAspArg 41
Db 193 AGCGGGGACATCGCTGCCGAG 252
QY 42 ArgGluProProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeu 61
Db 253 CAGGACCCGTCGAGACTGATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 312

Qy 62 GlnGlnLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArgProArgArg 81
Db 313 AACGACAGTGGCTGGCTGGAGAAAGTGCACCTTAAGGCGGAAACAGGGCTGAAAGA 372
Qy 82 AlaIleAlaAlaHisTyrGlyValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 101
Db 373 GCGATCGCAGCCATTATGAGTTCATCCAGCACTGGACAGAGCAGGCGAGCAGAGT 432
Qy 102 ValAspGlyThrValSerGlyTyrProGluLutThrLysIleAsnSerSerProLeuArg 121
Db 433 GTGGCGGAGACGTAGTGGCTGGAGGAGCAAGATCAACAGCTCAGCCCTCTGGCG 492
Qy 122 TyrAspArgGlnIleGlyLutPheThrValIleArgAlaGlyLeuTyrTyrLeuTyrCys 141
Db 493 TACAACCCCGCAGATGGGGAGTTTATGTCACCCGGGCTGGCTTACTACTGACTGCT 552
Qy 142 GlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeuValAsnGly 161
Db 553 CAGGTGCACCTTGTAGAGAGGGGAGGCTGTCTACTGAAGCTGGACTGCTGGTGATGT 612
Qy 162 ValIleAlaLeuArgCysLeuGlnGluPheSerAlaThrAlaIleAsnSerProGlyPro 181
Db 613 GTGCTGGCCCTGGCTGGCTGGAGAAATCTCAGCCACTGGCGGAGTTCCCTCGGGCC 672
Qy 182 GlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySerSerLeuArg 201
Db 673 CAGCTCGCCCTCTGTGCAGGTGTCTGGGCTGTGGCCCTGGCGGAGGCTCCCTCGCGG 732
Qy 202 IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 221
Db 733 ATCCGACCCCTCCCTGGGCGCCATCTCAAGGCTGCCCTTCCTCACTTCCGAGCTC 792
Qy 222 PheGlnValHis 225
Db 793 TTCCAGGTTTCAC 804
RESULT 6
AX180714 898 bp DNA Linear PAT 06-AUG-2001
LOCUS AX180714
DEFINITION Sequence 1 from Patent WO0145730.
ACCESSION AX180714
VERSION AX180714.1 GI:15132570
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 898)
AUTHORS Wiley,S.R.
TITLE Tweak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
IMMUNEX CORPORATION (US)
FEATURES
Source location/qualifiers
1..898
/organism="synthetic construct"
/db_xref="taxon:32630"
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=11
/protein_id="CAC50004.1"
/db_xref="gi:15132571"
/translation="MATGSRPSTSLDAFGLLCLPWLQEGSATSSDRMKQIEDKIEEILS
KIYHENEIARIKILIGERTSSLSRASISAOBPAOBEIYAEDODPSSEINPOTEE
ODPAPFLNRLVPRRSAPKGRKTRARAIANHYVHPRODGDGADCVGSGMEEA
RINSSPLRYRROJGERTVTRAGIYLYKCVHPEGKAVYIKIDLIDGVATLAKLEE
FSATPAASLSLGLQLKCOVSGLLAPRSSSLRKLRLPMAHLKAAPFLTYFGLEOVH"
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

Alignment Scores:
pred. No.: 1,15e-63
Score: 951.00
Percent Similarity: 93.24%

Length: 898
Matches: 184
Conservative: 9

Best Local Similarity: 88.89% Mismatches: 14
Query Match: 81.84% Indels: 0
DB: 6 Gaps: 0

US-09-245-198a-2 (1-225) x AX180714 (1-898)

Qy 19 SerLeuGlySerThrPalaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAla 38
Db 250 AGTTTGGGAGCCGGGCGCATCGCTGCCGCCAGAGCCTGCCAGAGAGAGCTGGTGCA 309
Qy 39 GlnAspArgArgGluProProGluLeuAsnProGlnThrGlnGluSerGlnAspVal 58
Db 310 GAGGAGAGCAGAGACCCCGGAACTGAATCCCGACAGAAAGAAAGCAGAGATCTCG 369
Qy 59 ProPheLeuGlnGluLeuValArgProArgArgSerAlaProLysGlyArgLysAlaArg 78
Db 370 CTTTTCGACCGCAGTACTAGTTCGGCTCGCAAGATGCACTTAAGGCGGAAACACGG 429
Qy 79 ProArgArgAlaIleAlaAlaHisTyrGlnValHisProArgProGlyGlnAspGlyAla 98
Db 430 GCTCGAAGACGATCGCAGCCCATTTATGAGTTCAACAGCACTGGACAGCAGGAGCG 489
Qy 99 GlnAlaGlyValAspGlyThrValSerGlyTyrProGluGluThrLysIleAsnSerSer 118
Db 490 CAGCGAGGTGGAGCGGAGAGTAGTGGCTGGAGAAAGCAAGATCAACAGCTCCAGC 549
Qy 119 ProLeuArgTyrAspArgGlnIleGlyLutPheThrValIleArgAlaGlyLeuTyr 138
Db 550 CCTTGGCCCTACAAACCGCAGATCGGGAGTTTATAGTACACCGGGGCTGGCTTACTAC 609
Qy 139 LeuTyrCysGlnValHisPheAspGlyGlyLysAlaValTyrLeuLysLeuAspLeu 158
Db 610 CTGTACTGTGACAGTCACTTTGATGAGGGAAAGCTGTCTACTCAAGCTGACTGCTG 669
Qy 159 ValAsnGlyValLeuAlaLeuArgCysLeuGluPheSerAlaThrAlaIleAsnSer 178
Db 670 GTGATGTGTGTGCTGGCCCTGGCTGGAGAAATCTCAGCCACTGGCGCGAGTTCC 729
Qy 179 ProGlyProGlnLeuArgLeuCysGlnValSerGlyLeuLeuProLeuArgProGlySer 198
Db 730 CTGGGGCCCGACCTCGGCTGGCTGGAGGTGTGGGCTGTGGCTGGCGGCGAGGCTCC 789
Qy 199 SerLeuArgIleArgThrLeuProThrPalaHisLeuLysAlaAlaProPheLeuThr 218
Db 790 TCCCTGGCGGATCCGACCCCTCCCGGGGCCATCTCAAGGCTGCCCTCTCTCACTAC 849
Qy 219 PheGlyLeuPheGlnValHis 225
Db 850 TTGGGACTCTTCCAGGTTTCAC 870
RESULT 7
BC019047 1651 bp mRNA Linear PRI 11-DEC-2001
LOCUS BC019047
DEFINITION Homo sapiens, similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.
ACCESSION BC019047
VERSION BC019047.1 GI:17512138
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1651)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (07-DEC-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-remail.nih.gov

Tissue Procurement: Louis Staudt
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/nisc_mgc@ngri.nih.gov
 Contact: Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dierich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stancil, P., Thomas, P.J., Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.lnl.gov>
 Series: IRAL Plate: 30 Row: P Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 4507596.

FEATURES

Location/Qualifiers
 1..1651
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="MGC:20669 IMAGE:4766071"
 /tissue_type="Primary B-Cells from Tonsils"
 /clone_id="NIH_MGC_48"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 106..510
 /codon_start=1
 /product="Similar to tumor necrosis factor (ligand)
 superfamily, member 12"
 /protein_id="AAH19047.1"
 /db_xref="gi:17512139"
 /translation="MAARRSQRRRGRRGEPGALLVPLALGSLALACILLLAVSL
 GSRASLSQEPQAEEDDPSLNLNPTGSDPAPFLNRLVPRRSAPKGRKTR
 ARRALAHEVHPRRGODGADQADGGYTTLRP"

CDS

BASE COUNT 344 a 517 c 481 g 309 t
 ORIGIN

Alignment Scores:
 Pred. No.: 5.76e-63 Length: 1651
 Score: 945.90 Matches: 198
 Percent Similarity: 68.09% Conservative: 9
 Best Local Similarity: 65.13% Mismatches: 17
 Query Match: 81.37% Indels: 81
 DB: Gaps: 1

US-09-245-198a-2 (1-225) x BC019047 (1-1651)

QY 2 LeuSerLeuGluLeuAlaLeuAlaCysLeuGluLeuLeuValValSerLeuGly 21
 DB 181 CTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCG 240
 QY 22 SerTrpAlaThrLeuSerAlaGlnGluProSerGlnGluLeuThrAlaGluAspArg 41
 DB 241 AGCGGGGCAACGCTGTCGCCGAGAGGCGTCCGAGAGAGAGTGGTGGCAGAGAGAG 300
 QY 42 ArgGluProGluLeuAsnProGlnThrGluGluSerGlnAspValValProPheLeu 61
 DB 301 CAGGAGCCGCGGAGACTGATCCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 360
 QY 62 GluGluLeuValArgProArgArgSerAlaProGlyGlyArgLysAlaArgProArgArg 81
 DB 361 AACCGACTAGTTCGGCTCCGACAGAGGACCTTAAGGCGGAAACACGCGCTCGAAGA 420
 QY 82 AlaAlaAlaAlaHisTrpGluValHisProArgProGlyGlnAspGlyAlaGlnAla 100
 DB 421 GCGATGCGACCCATATGATGATTCATCCAGCATGCGACGAGAGAGAGAGAGAGAG 479
 QY 100 ----- 100

DB 480 TGGAGGTACACAACTTGTCTGAGGCCATGAGATCTAAGTGTGGAGCCAAAGATTGAA 539
 QY 100 ----- 100
 DB 540 CCCAGCTAGATGTGCTGTGTACTGTGACATGTCTGCAATGAAGGCGAGTGTGTTG 599
 QY 100 ----- 100
 DB 600 CAGGGGTGAGGGGTCCATGACGAGGGCCACATCCAAAAGGAGAGGAAGTTCCAGAA 659
 QY 101 -----Gly 101
 DB 660 AGAAGAGAGACATCTCCACCATTTACCAGAGGCTCAAGAGAGAACAGACCAAGT 719
 QY 102 ValAspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerProLeuArg 121
 DB 720 GTGGAGGGGACACTGATGTGCTGGAGAGAACCAAGATCAACAGCTCCAGCTCTGCGC 779
 QY 122 TyrAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTyrrTyrrCys 141
 DB 780 TACACGCGCAGATCGGGAGGTTATAGTCACCCGGGCTGGCTTACTACTGTG 839
 QY 142 GlnValHisPheAspGluGlyLysAlaValTyrrLeuLysLeuAspLeuValAsnGly 161
 DB 840 CAGGTCCACTTGTATGAGGGAGGAGGCTGTCTACCTGAGCTGCTGTGTGATGTG 899
 QY 162 ValLeuAlaLeuArgCysGluGluGluGluPheSerAlaThrAlaAlaSerSerProGlyPro 181
 DB 900 GTCTGTGCGCTGGCTGCTGGAGAAATTCACGACACTGCGCGAGTTCCCTCGGGGCC 959
 QY 182 GlnLeuArgLeuCysGlnValSerGlyLeuAlaProLeuArgProGlySerSerLeuArg 201
 DB 960 CACCTCCGCGCTGCGCAGAGGTCTGGGCTGTGGCCCTCGGGCGCAGGTCCTCCGCG 1019
 QY 202 IleArgThrLeuProTrpAlaHisIleLeuLysAlaAlaProPheLeuThrTyrrPheGlyLeu 221
 DB 1020 ATCCGACCTCCCGCGGGCCATCAAGCGTCCCTCTCTCACTCACTTGAGATC 1079
 QY 222 PheGlnValHis 225
 DB 1080 TTCACAGTTTCAC 1091

RESULT 8
 AC069459 203083 bp DNA linear HTG 27-JUN-2001
 LOCUS Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT
 DEFINITION
 SEQUENCE 7 unordered pieces.
 ACCESSION AC069459 GI:14547768
 VERSION AC069459
 HTG: HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Mus musculus.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 203083)
 REFERENCE
 AUTHORS Metzker, M.L., Lewis, L.R., Hume, J., Edwards, C., Harris, C., Dederich, M.D., Thomas, S., Okwuon, G., Carllock, C., Garner, T., Addison, S., Pace, A., Williams, G., Bonini, D., Brooks, A., Brown, J., Bunay, C., Bunac, C., Burkett, C., Ding, Y., Dugan-Rocha, S., Fernandez, C., Ferraguto, D., Forcun-Tansey, J., Gill, R., Gorrell, J.H., Gunaratne, P., Haller, G., Hernandez, J., Hognes, M., Hosak, H., Hou, X., Huber, J., Jackson, L., Jia, Y., Kelly, J., Kelly, S., Kovar, C., Liu, J., Liu, W., Louised, H., Lozdo, R.J., Martin, R., Massey, E., McLeod, M.P., Mel, G., Moore, S., Morgan, M., Morris, S., Neal, D., Nelson, A., Nguyen, R., Nguyen, N., Ogih, M., Parish, B., Perez, L., Reiter, D., Say, J., Shen, H., Vasquez, L., Watlington, S., Williamson, A., Wrensford, G., Zhou, X., Bouck, J., Hodgson, A., Muzny, D.M., Rives, M., Scherer, S., Sodergren, E., Weinstock, G., Worley, K. and Gibbs, R.
 TITLE Direct Submission
 JOURNAL Unpublished

FEATURES This sequence is the entire insert of clone RP23-422L16.

source

1. 234182
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_lib="RPCT-23"
mlsc_feature complement(84050..84131)
note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
BASE COUNT 59310 a 56824 c 57519 g 60529 t
ORIGIN

Alignment Scores:

Pred. No.:	1 24e-34	Length:	234182
Score:	602.50	Matches:	125
Percent Similarity:	70.00%	Conservative:	1
Best Local Similarity:	69.44%	Mismatches:	0
Query Match:	51.85%	Indels:	54
	10	Gaps:	1

US-09-245-198a-2 (1-225) x AL603707 (1-234182)

QY 100 AlAGlyVAlAspJyThrVAlSerGlyTTPGluGluThrLysIleAsnSerSerPro 119
:::|||||
Db 66609 TCAGGTGTGATGGACAGAGTGCGCTGGAGAGACCAAAATCAACAGCTCCAGCCCT 66550
QY 120 LnuAGTytAspArgGlnIleGlyGluPhetVAlIleArgAlaGlyLeuTytTyleu 139
|||||
Db 66549 TCAGGTACAGACCCGACGATGGGAAATTACAGTCACGGCTGGCTCTACTACTG 66490
QY 140 TtCys----- 141
|||||
Db 66489 TACTGTAGGTAAAGCCCTGCTGCTCAGGGGACAGCAAGGCTAAGGGAGAGGA 66430
QY 141 ----- 141
Db 66429 CTGGCAAGAAATGGTGGAGAGGAGACACTCTGCACTAGAGAAAGTCCCTGCTG 66370
QY 142 -----GluValHisPheAs 146
Db 66369 TCATATGAGGAGGACAGAGCTGTGGATTTGCTTCCCTCTCTGTCCAGAGTGCATTGA 66310
QY 146 pGluGlyLysAlaValTytLeuLysLeuAspLeuValAsnGlyValLeuAlaLeuAr 166
Db 66309 TGAAGGAAAGAGCTGTCTACTGAAGGAGGACTTGTCTGTGACAGGCTGCTGGCCCTGCG 66250
QY 166 gCysLeuGluGluPheserAlaThrAlaIleSerSerProGlyProGlnLeuArgLeuCy 186
|||||
Db 66249 CTCCTGTGAAGAAATTCACACACAGCAAGCACTCTCCGCGCCACCTCGTTTGTG 66190
QY 186 sGlnValSerGlyLeuLeuProLeuArProGlySerSerLeuArgIleArgThrLeuPr 206
|||||
Db 66189 CCAAGTGTCTGGGCTGTGGCTGGCGGCAAGGCTCTCCCTTCGAGTCCGCAACCTCC 66130
QY 206 oTrrAlaHisLeuLysAlaIleArProPheLeuThrTytPheGlyLeuPheGlnValHis 225
|||||
Db 66129 CTGGGCTCATTTAAGGCTCCCTCTCTACTACTTGTGACACTTTTCAAGTTCAC 66072
RESULT 10
LOCUS AC016876/c 177703 bp DNA linear HTG 06-AUG-2002
DEFINITION Homo sapiens clone RP11-186B7, *** SEQUENCING IN PROGRESS ***, 10
unorderd pieces.
ACCESSION AC016876
VERSION AC016876
KEYWORDS HTG; HTGS_PHASE1; HTGS_FULFILL; HTGS_ACTIVATEIN.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 177703)

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

Birren, B., Nusbaum, C. and Lander, E.
Homo sapiens chromosome, clone RP11-186B7
Unpublished
2 (bases 1 to 177703)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, J., Bhat, N., Becker, R., Boguslavsky, L., Boukhalter, B.,
Brown, A., Castle, A., Collins, M., Collins, S., Collymore, A.,
Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., Johnson, R., Jones, C., Kann, L., Karats, A., Klein, J.,
Lehoczy, J., Liu, C., Locke, K., MacDonald, P., Margulis, N.,
McEwan, P., McGuire, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Testaye, S., Tittel, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W., J., Zimmer, A. and Zody, M.

TITLE
JOURNAL
REFERENCE
AUTHORS

Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 177703)

Birren, B., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A.,
Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., S., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kellis, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneses, L., Mihova, T., Mlewa, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C., H.,
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Testaye, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zemke, L., Zimmer, A. and Zody, M.

TITLE
JOURNAL
COMMENT

Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 2, 2002 this sequence version replaced g1:15421989.
All repeats were identified using RepeatMasker:
Smil, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu

Project Information
Center project name: L3849
Center clone name: 186_B_7

NOTE: This is a 'working draft' sequence. It currently
consists of 10 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 48645: contig of 48645 bp in length
* 48646 48745: gap of 100 bp
* 48746 48745: contig of 9532 bp in length
* 58278 58377: gap of 100 bp
* 58378 58377: contig of 7427 bp in length
* 58378 65804: contig of 7427 bp in length
* 65805 65904: gap of 100 bp
* 65905 79793: contig of 13889 bp in length
* 79794 79893: gap of 100 bp
* 79894 99493: contig of 19600 bp in length

```

* 99494 99593: gap of 100 bp
* 99594 111049: contig of 11456 bp in length
* 111050 111149: gap of 100 bp
* 111150 125020: contig of 13871 bp in length
* 125021 125120: gap of 100 bp
* 125121 145109: contig of 19989 bp in length
* 145110 145209: gap of 100 bp
* 145210 169458: contig of 24249 bp in length
* 169459 169558: gap of 100 bp
* 169559 177703: contig of 8145 bp in length.
FEATURES
    source
        1..177703
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone="RP11-186B7"
        /clone_11b="RP11-11 Human Male BAC"
BASE COUNT  42862 a 45967 c 46568 g 41254 t 1052 others
ORIGIN
Alignment Scores:
Pred. No.:      9,71e-31      Length:      177703
Score:          549.50        Matches:      117
Percent Similarity: 63.02%    Conservative: 4
Best Local Similarity: 60.94%  Mismatches:  5
Query Match:     47.29%      Indels:       66
DB:              2           Gaps:          1

US-09-245-198a-2 (1-225) x AC016876 (1-177703)
OY  100 Alaglyvalaspjlythrvalserglytrpglugluhrlylsleasnserserperro 119
    ::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::::
DB  37636 TCAGGTGTGACGAGGAGAGTGTGCTGGTGGAGGAGCAATCAACAGCTCCAGCCCT 37577
OY  120 leuargtyrAsparglnllelygluphethrvallleargAlaglyleutyrryleu 139
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB  37576 CTGGCTTACACACCGCCAGATCGGGAGTTATATGTCACCCGGGCTGGCTTACTACCTG 37517
OY  140 TTYCYS-----
    |||||
DB  37516 TACTGTCAAGTAAAGCCCACTGTGGCTGCATGGGTAAAGCAGTAAAGAGTGGCGAAGGT 37457
OY  141 -----
    |||||
DB  37456 TTGCCAGAGAGTGGGAGACAGTACAGGCTGGAGGAGGTGAGTTGGGTGGG 37397
OY  141 -----
    |||||
DB  37396 ATGGGATGCTCTGCTGCTGAGAAATTGGAATTGAGCGAGGCGAGCAGAGGCTTGG 37337
OY  142 -----
    |||||
DB  37336 ACTGGCCCTGTTGTCCCAACCCAGGTGACATTGTGAGGGGAGAGGCTGTCTACTGAA 37277
OY  154 sleuaspLeuLeuValasnGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaTh 174
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB  37276 GCTGGACTTGCTGTGTGAGATGATGTGCTGCTGCGCTCCCTGGAGGATTTCTCAGCCAC 37217
OY  174 TAlaAlaSerSerProGlyProGlnLeuArgLeucysGlnValSerGlyLeuLeuProle 194
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB  37216 TGGCGCGAGTTCTCCCTGGGCCCCACCTCGCTCGCTGCCAGGTGTGCTTGGCCCT 37157
OY  194 uArgProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisleuLyAlaAlaP 214
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB  37156 GCGGCGAGGCTCTCCCTCGGATCCGACCTCCCTGGGCGCAATCTCAAGGCTGCCCC 37097
OY  214 oPheLeuThrTyrrPheGlyLeuPheGlnValHis 225
    |||||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
DB  37096 CTTCCTCACCTACTTCGAGCTCTCCAGGTTCAC 37063
RESULT 11
LOCUS      AC127470          215795 bp      DNA      linear      HTG 17-JUL-2002
DEFINITION Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15

```

```

unordered pieces.
AC127470
AC127470.1 GI:21886866
HTG: HTGS_PHASE1: HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Pan.
1 (bases 1 to 215795)
REFERENCE
AUTHORS
Akhter,N., Antonellis,A., Ayale,K., Beckstrom-Stenberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Latic,P.,
Lee-Lin,S.-O., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Margulies,E.H., Masciello,C., Maskeri,B., Mastrian,S.D.,
McCluskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schneider,M.G., Stantirpop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 215795)
Green,E.D.
Direct Submission
Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoehgri.nih.gov
----- Project Information
Center project name: cms
Center phone name: 145D13
----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 203929 bases at least Q40
Consensus quality: 206865 bases at least Q30
Consensus quality: 209181 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 214395; sum-of-ctnigs
Quality coverage: 7.30x in Q20 bases; agarose-fp
Quality coverage: 6.47x in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

```

1 3593: contig of 3593 bp in length
* 3594 3693: gap of unknown length
* 3694 8001: contig of 4308 bp in length
* 8002 8101: gap of unknown length
* 8102 15639: contig of 7538 bp in length
* 15640 15739: gap of unknown length
* 15740 24775: contig of 9036 bp in length
* 24776 24875: gap of unknown length
* 24876 30150: contig of 5275 bp in length
* 30151 30250: gap of unknown length
* 30251 39212: contig of 8962 bp in length
* 39213 39312: gap of unknown length
* 39313 49342: contig of 10030 bp in length
* 49343 49442: gap of unknown length
* 49443 60956: contig of 11514 bp in length
* 60957 61057: gap of unknown length
* 61057 74520: contig of 13464 bp in length
* 74521 74621: gap of unknown length
* 74621 89522: contig of 14902 bp in length

```


AUTHORS Akhter, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C.,
Brooks, S., Dietrich, N.L., Grante, S., Guan, X., Gupta, J.,
Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Lalic, P.,
Lee, L.N., S.O., Legaspi, R., Maduro, O.L., Maduro, Y.B.,
Margulies, E.H., Mastello, C., Maskeri, B., Mascian, S.D.,
McCluskey, J.C., McDowell, J., Peguigian, C., Pearson, R.,
Portnoy, M.E., Prasad, A., Schueler, M.G., Stantrop, S., Thomas, J.W.,
Thomas, P.J., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

TITLE NISC Comparative Sequencing Initiative

JOURNAL Unpublished

REFERENCE 2 (pages 1 to 161428)

AUTHORS Green, E.D.

JOURNAL Direct Submission

JOURNAL Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
Grovefront Circle, Gaithersburg, MD 20877, USA

COMMENT ----- Genome Center

Center: NIH Intramural Sequencing Center

Center code: NISC

Web site: <http://www.nisc.nih.gov>

Contact: nisc.zoo@nih.gov

----- Project Information

Center project name: cwp

Center clone name: 332E11

----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads

Chemistry: dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.990319

Consensus quality: 148712 bases at least Q40

Consensus quality: 151471 bases at least Q30

Consensus quality: 153125 bases at least Q20

Insert size: 152000; agarose-fp

Insert size: 160028; sum-of-contigs

Quality coverage: 5.36x in Q20 bases; agarose-fp

Quality coverage: 5.10x in Q20 bases; sum-of-contigs

----- NOTE: This is a 'working draft' sequence. It currently
* consists of 15 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 4734: contig of 4734 bp in length

4735 4834: gap of unknown length

4835 7325: contig of 2491 bp in length

7326 7425: gap of unknown length

7426 1261: contig of 4836 bp in length

1262 1261: contig of unknown length

1262 19031: contig of 6670 bp in length

19032 19031: gap of unknown length

19032 27849: contig of 8718 bp in length

19132 27849: gap of unknown length

27850 31801: contig of 3852 bp in length

31802 31901: gap of unknown length

31902 40781: contig of 8880 bp in length

40782 40881: gap of unknown length

40882 49783: contig of 8902 bp in length

49784 49883: gap of unknown length

49884 59150: contig of 9267 bp in length

59151 59250: gap of unknown length

59251 68679: contig of 9429 bp in length

68680 68779: gap of unknown length

68780 78984: contig of 10205 bp in length

78985 79084: gap of unknown length

79085 93446: contig of 14362 bp in length

93447 93547: gap of unknown length

102176 102175: contig of 8629 bp in length

102276 102275: gap of unknown length

13185 13185: contig of 29410 bp in length

13186 13185: gap of unknown length

FEATURES * 131786 161428: contig of 29643 bp in length.

source Location/Qualifiers

1. 161428

/organism="Canis familiaris"

/db_xref="taxon:9615"

/clone="RP81-332E11"

/clone_lib="RP81"

1. 4734

/note="assembly-fragment"

4835. 7325

/note="assembly-fragment"

7426. 12261

/note="assembly-fragment"

12362. 19031

/note="assembly-fragment"

19132. 27849

/note="assembly-fragment"

27950. 31801

/note="assembly-fragment"

31902. 40781

/note="assembly-fragment"

40882. 49783

/note="assembly-fragment"

49884. 59150

/note="assembly-fragment"

59251. 68679

/note="assembly-fragment"

68780. 78984

/note="assembly-fragment"

79085. 93446

/note="assembly-fragment"

93547. 102175

/note="assembly-fragment"

clone_end:SP6

vector_side:right"

102276. 131685

/note="assembly-fragment"

131786. 161428

/note="assembly-fragment"

BASE COUNT 39586 a 41863 c 41233 g 37340 t 1406 others

ORIGIN

Alignment Scores:

Pred. No.: 2.1e-30 Length: 161428

Score: 544.50 Matches: 116

Percent Similarity: 63.35% Conservative: 5

Best Local Similarity: 60.73% Mismatches: 12

Query Match: 46.86% Indels: 58

DB: 2 Gaps: 1

US-09-245-198a-2 (1-225) x AC126925 (1-161428)

Oy 93 ProGlylnmspglylaGlnAlaGlyValAspGlyThrValserGlyTrpGluGluThr 112

Db 42806 CCAATCTCTGAACCCCTGCATTCAAGGTGAGTGGAGGTGAGTGGAGAGGCC 42747

Oy 113 LysIleasnserSerSerProLeuArgTrpAspArgGlnIleGlyGluPheThrValIle 132

Db 42746 AAATATCAAGCTCCAAACCCACGCGGTATGACCGCAGGCGGGAATTTAATGTCACG 42687

Oy 133 ArgAlaGlyLeuTrpTrpTrpLeuTrpCys----- 141

Db 42686 CGGGCTGCGCTTACCTGCTACGTGACGTAAGCCCACTGCTCCACGGTAGGCA 42627

Oy 141 ----- 141

Db 42626 GGAAGCGGGGCCCCAGGTTAGAGGAGACGAGAGAGGTTTGTTCAGGGGGCAGTTGGC 42567

Oy 141 ----- 141

Db 42566 GTGGGAGCCTTAGGCTGGGGGCCGATGGATCTCTGGGGTCACCAAGAGCCCGGACTCT 42507

Oy 142 -----GlnValHisPheAspGluGlyLysAlaValTrpLeuLysIle 155

```

Db 42506 GGGGCTCTCGCGGCCCCAGGTGACTTGTATGAGGGAGGAGCTGTACTGTAGAGCT 42447
Qy 155 uaspleuteuValasnglyValaLeuArycysleugluPheseraAlaThral 175
Db 42446 GGACTTGGTGGTATGAGCGCCCTGGCTGCTCCGTAAGAGTTCTCCGCCACAGC 42387
Qy 175 alaserseProglyProgluLeuArycysglnValaserglyLeuLeuProleuAr 195
Db 42386 GCGCAGCAGCTGGGCGCCAGCTCGCTGCGCAAGTGTGGGCTTTCGCCCTTCGG 42327
Qy 195 gProglySerseLeuArygileAryThreupProTrAlaHisleuAlaAlaProph 215
Db 42326 GCGCGGGTCTCTCCGCGATCGGACCCCTCCCTGGGCCCATTCAGAGCGCCGCCCTT 42267
Qy 215 eLeuthrTyrrhPheglyLeuPheglnValHis 225
Db 42266 CCTACTACTTCTGCGACTCTCTCAGGTTAAC 42236

RESULT 14
AC119115 138792 bp DNA linear HTG 18-JUL-2002
DEFINITION Rattus norvegicus clone CH230-320N23, *** SEQUENCING IN PROGRESS
ACCESSION AC119115.2 GI:21746718
VERSION HTG; HTGS_PHASE1.
KEYWORDS
SOURCE Norway rat.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
1 (bases 1 to 138792)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Alsbrooks,S.L., Amaralunga,H.C., Are,J.R., Ayele,M., Banks,T.,
Barbarta,J., Benton,J., Blinage,K., Blankenburg,K., Bonin,D.,
Bouck,J., Bowle,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Dayla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
DeLaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Doutwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,U., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Haylak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karstson,E., Kelly,S., Khan,U., King,L., Korvan,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Licharge,O., Lieu,C., Liu,J., Liu,W., Louisedge,H.,
Lozdo,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mathney,E., McLeod,M.P., Meador,M., Mel,G., Metzker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Mosser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenwu,S., Ogun,M., Okunolu,G.,
Orangunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peterson,L., Pickens,R., Primus,E., Pu,L., Quiles,M., Ren,Y.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G.,
Scherer,S., Scott,G., Shen,H., Shoshitari,N., Sisson,I.,
Sodergren,E., Sonalike,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Syatke,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Umanli,K., Vasquez,L., Vera,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wlarczyk,R., Wooden,S., Wolley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Direct Submission
Unpublished
2 (bases 1 to 138792)
Worley,K.C.
Direct Submission
Submitted (25-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 138792)
Worley,K.C.
Direct Submission
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20303440.
----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
----- Project Information
Center project name: GUXG
Center clone name: CH230-320N23
----- Summary Statistics
Sequencing vector: Plasmid
Chemistry: Dye-terminator Big Dye 1.00% of reads
Assembly program: Phrap, version 0.990329
Consensus quality: 112720 bases at least Q40
Consensus quality: 11666 bases at least Q30
Consensus quality: 119165 bases at least Q20
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a "working draft" sequence. It currently
* consists of 32 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1023: contig of 1022 bp in length
1123: gap of unknown length
1123: contig of 1272 bp in length
2494: gap of unknown length
2495: gap of unknown length
3663: contig of 1169 bp in length
3664: gap of unknown length
3763: gap of unknown length
4879: contig of 1116 bp in length
4880: gap of unknown length
4980: gap of unknown length
6787: contig of 1808 bp in length
6788: gap of unknown length
6887: gap of unknown length
8192: contig of 1305 bp in length
8193: gap of unknown length
8292: gap of unknown length
8293: contig of 1169 bp in length
9461: gap of unknown length
9462: gap of unknown length
9562: contig of 1874 bp in length
11435: contig of 1874 bp in length
11436: gap of unknown length
11436: contig of 2000 bp in length
11536: gap of unknown length
11536: contig of 2000 bp in length
13536: gap of unknown length
13536: contig of 3688 bp in length
17324: gap of unknown length
17324: contig of 2993 bp in length
20406: contig of 2993 bp in length
20407: gap of unknown length
20407: contig of 2993 bp in length
23445: gap of unknown length
23445: contig of 2993 bp in length
23446: gap of unknown length
23446: contig of 2055 bp in length
25600: gap of unknown length
25601: gap of unknown length
25701: contig of 2349 bp in length
28049: gap of unknown length
28050: contig of 2325 bp in length
30474: contig of 2325 bp in length
30475: gap of unknown length
30574: gap of unknown length
30575: contig of 4924 bp in length
35598: gap of unknown length

```

```

* 35599 39284: contig of 3686 bp in length
* 39285 39384: gap of unknown length
* 39385 41970: contig of 2586 bp in length
* 41971 42070: gap of unknown length
* 42071 45659: contig of 3589 bp in length
* 45660 45759: gap of unknown length
* 45760 50332: contig of 4573 bp in length
* 50333 50432: gap of unknown length
* 50433 54695: contig of 4263 bp in length
* 54696 58289: gap of unknown length
* 58290 58390: contig of 3494 bp in length
* 58391 63599: gap of unknown length
* 63600 63699: gap of unknown length
* 63700 70305: contig of 6606 bp in length
* 70306 76123: gap of unknown length
* 76124 76223: gap of unknown length
* 76224 84961: contig of 8738 bp in length
* 84962 93614: gap of unknown length
* 93615 93714: gap of unknown length
* 93715 103352: contig of 9638 bp in length
* 103353 103452: gap of unknown length
* 103453 110299: contig of 6847 bp in length
* 110300 110399: gap of unknown length
* 110400 116180: contig of 5781 bp in length
* 116181 116280: gap of unknown length
* 116281 125484: contig of 9204 bp in length
* 125485 125584: gap of unknown length
* 125585 138792: contig of 13208 bp in length.
Location/Qualifiers
1. 138792
/organism="Rattus norvegicus"
/db_xref="taxon:10116"
/clone="CH230-320N23"

```

```

BASE COUNT 34391 a 31996 c 31712 g 36193 t 4500 others
ORIGIN

```

```

Alignment Scores:
Pred. No.: 7 23e-30 Length: 138792
Score: 536.50 Matches: 110
Percent Similarity: 67.46% Conservative: 2
Best Local Similarity: 65.09% Mismatches: 4
Query Match: 46.17% Indels: 53
DB: 2 Gaps: 1

```

```

US-09-245-198a-2 (1-225) x AC119115 (1-138792)

```

```

QY 101 GlyValAspGlyThrValSerGlyTrpGluGluThrLysIleAsnSerSerProLeu 120
|||||
Db 137345 GGTCGTGATGGAGACAGTGAAGGCTGGAGAGACCAAAATCAACACCTCCAGCCCTCTG 137404
QY 121 ArgGlyTrpAspArgGlnIleGlyGluPheThrValIleArgAlaGlyLeuTYrTYrLeuTYr 140
|||||
Db 137405 CGCATATACCGCCAGATTGGGGAATTACGGTCATCAGGCGCTGACTACTACTCTGAC 137464
QY 141 Cys----- 141
|||
Db 137465 TGTCAAGTAAAGCCCTGGGCTCCATGGGTAGACGATGGCTAAGGGAGAAAGCTTGCCA 137524
QY 141 ----- 141
Db 137525 AGAATATGGTGGAGTGGAGAACCTGGGTTTCATGAAGAGATGCTGGATTTCGATG 137584
QY 142 -----GlnValHisPheAspGlu 147
|||||
Db 137585 AGGAGCAGCAGAGAGTCTGATTTTCTGCTCTCTCTCTCCAGCTGACATTGATGAG 137644
QY 148 G1yValAlaValTyrLeuLysLeuAspLeuValAsnGlyValIleuAlaLeuArgCys 167
|||||
Db 137645 GGGAAAGCAGTCTTACTCTGAAGCTGGACTGCTGATGATGTGTGCTGGCCCTGCGCTGC 137704

```

```

QY 168 LeuGluGluPheSerAlaThrAlaAlaSerSerProGlyProGluLeuArgCysGln 187
|||||
Db 137705 CTGGAGAAATTCACGACAGCAGAGCTCTCTCTGCGCCAGCTCTCTGTCACG 137764
QY 188 ValSerIleuLeuPheProLeuArgProGlySerSerIleuArgIleArgThrLeuProTrp 207
|||||
Db 137765 GTGCTGGGCTGTGGCTTCGCGACAGGGGTCTTCCCTTGATCGCTACCATCACTTGC 137824
QY 208 AlaHisLeuLysAlaAlaProPheLeu 216
|||||
Db 137825 AGTCATATTAAAGCGGCTACCTTCGTT 137851

```

```

RESULT 15
AC126921 153553 bp DNA linear HTG 10-JUL-2002
LOCUS Bos taurus clone RP42-45024, WORKING DRAFT SEQUENCE, 12 unordered
DEFINITION
AC126921
AC126921.1 GI:21724098
VERSION
KEYWORDS
SOURCE
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 153553)
REFERENCE
AUTHORS
Ahkter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Grantham,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Laric,P.,
Lee-Plin,S.-O., Legaspi,R., Maduro,O.L., Maduro,V.B.,
Margulies,E.H., Mastello,C., Maskerli,B., Masrihan,S.D.,
McCloskey,J.C., McDowell,J., Pegurigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Shanripop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgren,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 153553)
Green,E.D.
Direct Submission
Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
GroveMont Circle, Galtersburg, MD 20877, USA
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoem@nih.gov
----- Project Information
Center clone name: 045D24

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT

```

```

----- Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 146066 bases at least Q40
Consensus quality: 147748 bases at least Q30
Consensus quality: 148824 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 152453; sum-of-ctrls
Quality coverage: 8.80x in Q20 bases; agarose-fp
Quality coverage: 8.72x in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2312: contig of 2312 bp in length
* 2313 2412: gap of unknown length

```

```

*      2413      5841: contig of 3429 bp in length
*      5842      5941: gap of unknown length
*      5942      8435: contig of 2494 bp in length
*      8436      8535: gap of unknown length
*      8536      15799: contig of 7264 bp in length
*      15800      15899: gap of unknown length
*      15900      25224: contig of 9325 bp in length
*      25225      25324: gap of unknown length
*      25325      32504: contig of 7180 bp in length
*      32505      32604: gap of unknown length
*      32605      40970: contig of 8366 bp in length
*      40971      41070: gap of unknown length
*      41071      56590: contig of 15520 bp in length
*      56591      56690: gap of unknown length
*      56691      73770: contig of 17079 bp in length
*      73770      73869: gap of unknown length
*      73870      90859: contig of 16990 bp in length
*      90860      90959: gap of unknown length
*      90960      111428: contig of 20469 bp in length
*      111429      111528: gap of unknown length
*      111529      153553: contig of 42025 bp in length.
FEATURES
Source
1..153553
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="RP42-45D24"
/clone_lib="RP42"
1..2312
/note="assembly-fragment"
2413..5841
/note="assembly-fragment"
clone_end:r7
vector_side:left"
5942..8435
/note="assembly-fragment"
8536..15799
/note="assembly-fragment"
15900..25224
/note="assembly-fragment"
25325..32504
/note="assembly-fragment"
32605..40970
/note="assembly-fragment"
41071..56590
/note="assembly-fragment"
56691..73770
/note="assembly-fragment"
73870..90859
/note="assembly-fragment"
90960..111428
/note="assembly-fragment"
111529..153553
/note="assembly-fragment"
clone_end:SP6
vector_side:right"

```

```

BASE COUNT      39178 a      37393 c      36616 g      39259 t      1107 others
ORIGIN

```

```

Alignment Scores:
Pred. No.:      2,81e-27      Length:      153553
Score:      503.00      Matches:      111
Percent Similarity:      58.59%      Conservative:      5
Best Local Similarity:      56.06%      Mismatches:      17
Query Match:      43.29%      Indels:      66
DB:      2      Gaps:      1

```

```

US-09-245-198a-2 (1-225) x AC126921 (1-153553)

```

```

QY      93      ProGluGlnAspGlyAlaGlnIaGlyValAspGlyThrValSerGlyTrpGluGlnThr 112
DB      26848      CCTGCTCTGAAACACCCACATTCAGGTGTGACGCGGTGAGTGGCGAGAGGCC 26907
QY      113      LysIleAsnSerSerProLeuArgTyrAspArgGlnIleGlyGluPheThrValIle 132

```

```

DB      26908      AAATATACAGCTCCCAACCCCTGGCTATGACTGCGACACCGGGCAATTTCAGTCAC 26967
QY      133      ArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal----- 143
DB      26968      CGGGCTGGCTGTACTACTGTAAGCTGTAGGT- AAGCCCACTGGCTCCAGGGTAAG 27026
QY      143      ----- 143
DB      27027      CCGGAACGTAGAGAGAAAGGCTGGGCTTCGGGGTTGGGGCAAGTTAAAGTGGGAGGG 27086
QY      143      ----- 143
DB      27087      GAGCGTGGGTTTGGGCTGAGAGGAGCCTTGGGCTCTAAGAGACACTGAGATGAAGCCC 27146
QY      144      -----HisPheAspGlu 147
DB      27147      AGGGCCAGCAGAGGCTGGACTCCGCCCTCCCTGCCCCAGGTGCACCTTGATGAG 27206
QY      148      GlyLysAlaValTyrLeuLysIleuAspLeuValAsnGlyValIleuAlaLeuArgCys 167
DB      27207      GGGAGGCTGTCTACTGTAAGCTGGACTGTGTGTGATGACACGCTGGCCCTGGCTGC 27266
QY      168      LeuGluGluPheSerAlaThrAlaAlaSerProGlyProGlnLeuArgLeuCysGln 187
DB      27267      CTGGAGGAATTCCTCGGCCACTGGGCGGCGAGTCCCTGGGCGCCAGCTCCGTCTGCCAA 27326
QY      188      ValSerGlyLeuLeuProLeuArgProGlySerSerLeuArgIleArgThrLeuProTrp 207
DB      27327      GTGTCAAGGCTCTTGTGGCCTGGCGGCGAGGCTCTTCCCTCGGATCCGACCCCTCCAG 27386
QY      208      AlaHisLeuLysAlaAlaIleProPheLeuThrTyrPheGlyLeuPheGlnValHis 225
DB      27387      ACCCACTCAAGGCTCCCCCTTCTCCTCACTTGTGACTCTTCCAAATTTCAC 27440

```

```

Search completed: March 31, 2003, 04:35:32
Job time : 3276.48 secs

```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.P5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 05:28:45 ; Search time 42.4361 Seconds
(without alignments)
706.506 Million cell updates/sec

Title: US-09-245-198a-2

Perfect score: 1162
Sequence: 1 VLSTGLALACGLLVVSL.....PMAHLKAPFLTYRGLRQVH 225

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1983.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1984.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1985.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1986.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1987.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1988.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1989.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1990.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1991.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1992.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1993.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1994.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1995.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1996.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1997.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1998.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA1999.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2000.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2001.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqp-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1162	100.0	225	19	AAW47524
2	1162	100.0	225	19	AAW47524
3	1089	93.7	211	20	AAW3591
4	1020	87.8	249	20	AAW29745
5	1020	87.8	249	20	AAW09369
6	1020	87.8	249	21	AAW07526
7	1020	87.8	249	21	AAW5338
8	1020	87.8	249	22	AAW00891
9	1020	87.8	249	23	AAW06129
10	1020	87.8	284	19	AAW47525

11	951	81.8	208	20	AAW3590
12	951	81.8	273	22	AAU03499
13	742	63.9	189	19	AAW29746
14	742	63.9	189	22	AAW00892
15	721	62.0	146	22	AAE00895
16	111	9.6	406	23	AAU77717
17	110.5	9.5	409	23	AAU77718
18	107.5	9.3	325	22	ABW67553
19	106.5	9.2	294	19	AAW69356
20	106.5	9.2	294	19	AAW68292
21	106.5	9.2	294	22	AAE08737
22	106.5	9.2	294	22	AAE04425
23	106.5	9.2	294	22	AAE01992
24	103	8.9	220	22	AAW62340
25	102.5	8.8	316	19	AAW83194
26	102.5	8.8	316	19	AAW83017
27	102.5	8.8	316	19	AAW59654
28	102.5	8.8	316	20	AAV17874
29	102.5	8.8	316	21	AAV91024
30	102.5	8.8	316	21	AAV84418
31	102.5	8.8	316	21	AAV84419
32	102.5	8.8	316	23	AAU78289
33	99	8.5	234	22	AAW62339
34	97	8.3	255	22	AAW6311
35	96.5	8.3	409	23	AAU77716
36	92	7.9	261	23	AAW49225
37	91	7.8	182	18	AAW08128
38	91	7.8	294	18	AAW09123
39	90.5	7.8	211	21	AAV58216
40	90.5	7.8	260	21	AAV58215
41	90.5	7.8	318	22	AAW82092
42	90.5	7.8	531	10	AAU91776
43	90	7.7	182	18	AAW24011
44	90	7.7	182	18	AAW24013
45	90	7.7	261	18	AAW09115

ALIGNMENTS

RESULT 1	
AAW47524	AAW47524 standard; Protein: 225 AA.
ID	AAW47524 standard; Protein: 225 AA.
AC	AAW47524;
DT	21-JUL-1998 (first entry)
XX	
DE	Mus musculus tumour necrosis factor related ligand (TREL).
XX	
DE	TREL: tumour necrosis factor related ligand; tnfr: treatment; cancer: autoimmune disease; immune system: stimulation; suppression; graft rejection.
KW	
XX	
OS	Mus musculus.
XX	
FT	Key
FT	Domain
XX	
XX	Location/Qualifiers
XX	1..21
XX	/note= "hydrophobic, transmembrane domain"
XX	
XX	W09805783-A1.
XX	
XX	12-FEB-1998.
XX	
XX	07-AUG-1997;
XX	97WO-US13945.
XX	
XX	18-MAR-1997;
XX	97US-0040820.
XX	07-AUG-1996;
XX	96US-0023541.
XX	18-OCT-1996;
XX	96US-0028515.
XX	
XX	(BIOT) BIOGEN INC.
XX	(UYGE-) UNIV GENEVA FACULTY MEDICINE.

Human TNRL3 protel
TWEAK extracellular
TNF related endoth
Human ULA1lag TREP
Human TREPA (TNF r
Drosophila melanog
Drosophila melanog
Drosophila melanog
NF-kB receptor act
NF-kB receptor act
Murine receptor ac
Murine receptor ac
Gp120 V3 loop-CD15
Human osteoprotege
Osteoclastogenesis
Amino acid sequenc
Murine TRANCE. Mu
Mouse OBW protein
Amino acid sequenc
Amino acid sequenc
Mouse TRANCE prote
Gp120 V3 loop-CD15
Fusion construct C
Drosophila melanog
Human CD145 mutant
CD40 ligand/zipper
CD40 ligand/zipper
Canine mature CD15
Canine CD154. Can
Rat osteoclast dif
Germ cell alkaline
CD40 ligand/zipper
CD40 ligand/zipper
Human CD40L mutein

PI Brownling JL, Chicheportliche Y;
 XX
 DR WPI: 1998-145619/13.
 DR N-PSDB; AAV18599.
 XX
 PT Tumour necrosis factor related ligand - useful for, e.g. treating
 PT cancer, auto-immune disease and immune responses to tissue grafts
 XX
 PS Claim 12; Pages 48-50; 69pp; English.
 XX
 CC The sequence is that of mouse tumour necrosis factor related
 CC ligand (TRELL). TRELL or active fragments can be included with a
 CC carrier in pharmaceutical compositions to treat cancer, auto-immune
 CC diseases or immune responses to tissue grafts, or to stimulate or
 CC suppress the immune system. It is useful to screen for TRELL
 CC receptors, by labelling with a detectable label and screening
 CC compositions for binding. Agents interfering with TRELL-receptor
 CC binding can also be screened for, can then be administered,
 CC optionally with interferon- gamma, to induce cell death or
 CC treat, suppress or alter immune responses (especially involving human
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its
 CC receptor. It's coding sequence can be used in gene therapy for
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
 CC auto-immune and inflammatory diseases or inherited genetic disorders,
 CC by introducing into cells, and expressing, therapeutically effective
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
 CC It may also be of use in the preparation of probe probes for
 CC screening natural/synthetic DNAs for TRELL-encoding sequences
 CC and for antisense therapy.
 CC
 XX
 SQ Sequence 225 AA:
 Query Match 100.0%; Score 1162; DB 19; Length 225;
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLSIGLALACGLLVVSGSWATLSAQPSEELTAEDRREPPELNPQTESQDVVPF 60
 DB 1 VLSIGLALACGLLVVSGSWATLSAQPSEELTAEDRREPPELNPQTESQDVVPF 60
 QY 61 LEQVLRPRRSAPKGRKARPRRAITAHYEVHPRGQDAQAGVDGTSGWETKINSSSPL 120
 DB 61 LEQVLRPRRSAPKGRKARPRRAITAHYEVHPRGQDAQAGVDGTSGWETKINSSSPL 120
 QY 121 RYDQIQIEFTVIRAGLYLYCQVHFDEGKAVYLLKDLVNGVIALRCLEEFSAATASSPG 180
 DB 121 RYDQIQIEFTVIRAGLYLYCQVHFDEGKAVYLLKDLVNGVIALRCLEEFSAATASSPG 180
 QY 181 POLRLCOVSGLLPRPGSSLRIRTLPAWHLKAAPFLTYFGLFOVH 225
 DB 181 POLRLCOVSGLLPRPGSSLRIRTLPAWHLKAAPFLTYFGLFOVH 225
 RESULT 2
 AAB07527
 ID AAB07527 standard; protein; 225 AA.
 XX
 AC AAB07527;
 XX
 DT 20-OCT-2000 (first entry)
 XX
 DE Amino acid sequence of a soluble recombinant murine TWEAK protein.
 XX
 KW TWEAK protein; immunological disorder; immune response; inflammation;
 KM TWEAK blocking agent; auto-immune disease; organ transplant rejection;
 KM Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour.
 XX
 OS Mus sp.
 XX
 PN WO200042073-A1.
 XX
 PD 20-JUL-2000.

PF 14-JAN-2000; 2000WO-0501044.
 XX
 XX 15-JAN-1999; 99US-0116168.
 XX
 PA (BIO) BIOGEN INC.
 XX
 PI Renner P;
 XX
 DR WPI: 2000-476036/41.
 XX
 PT Preventing and treating immune responses using modulators, especially
 PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for
 PT treating e.g. inflammation and graft versus host disease.
 XX
 PS Disclosure; Fig 1; 45pp; English.
 XX
 CC The present sequence represents a TWEAK protein. The specification
 CC describes a method for preventing or treating an immunological
 CC disorder and/or inhibiting an immune response in an animal. The
 CC method comprises administering a TWEAK blocking agent. The method may
 CC be used for preventing and treating immune disorders associated with
 CC inappropriate expression and/or activity of TWEAK. These disorders
 CC include autoimmune diseases, acute and chronic inflammation, organ
 CC transplant rejection, Graft-versus-host disease (GVHD), lymphoid cell
 CC malignancies, septic and other forms of shock, loss of immune
 CC responsiveness (as seen in human immunodeficiency virus (HIV)
 CC infections) and failure of the immune response to tumour growth.
 CC
 XX
 SQ Sequence 225 AA:
 Query Match 100.0%; Score 1162; DB 21; Length 225;
 Best Local Similarity 100.0%; Pred. No. 1.3e-112;
 Matches 225; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 VLSIGLALACGLLVVSGSWATLSAQPSEELTAEDRREPPELNPQTESQDVVPF 60
 DB 1 VLSIGLALACGLLVVSGSWATLSAQPSEELTAEDRREPPELNPQTESQDVVPF 60
 QY 61 LEQVLRPRRSAPKGRKARPRRAITAHYEVHPRGQDAQAGVDGTSGWETKINSSSPL 120
 DB 61 LEQVLRPRRSAPKGRKARPRRAITAHYEVHPRGQDAQAGVDGTSGWETKINSSSPL 120
 QY 121 RYDQIQIEFTVIRAGLYLYCQVHFDEGKAVYLLKDLVNGVIALRCLEEFSAATASSPG 180
 DB 121 RYDQIQIEFTVIRAGLYLYCQVHFDEGKAVYLLKDLVNGVIALRCLEEFSAATASSPG 180
 QY 181 POLRLCOVSGLLPRPGSSLRIRTLPAWHLKAAPFLTYFGLFOVH 225
 DB 181 POLRLCOVSGLLPRPGSSLRIRTLPAWHLKAAPFLTYFGLFOVH 225
 RESULT 3
 AAW93591
 ID AAW93591 standard; protein; 211 AA.
 XX
 AC AAW93591;
 XX
 DT 18-JUN-1999 (first entry)
 XX
 DE Mouse TNRL3 protein.
 XX
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KM developmental abnormality; gestational abnormality; prostate cancer;
 KM APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KM apoptosis; mouse.
 XX
 OS Mus sp.
 XX
 PN WO9911791-A2.
 XX
 PD 11-MAR-1999.

PF 04-SEP-1998; 98WO-US18393.
XX
PR 05-SEP-1997; 97US-0924634.
XX
PA (UNITW) UNIV WASHINGTON.
XX
PI Chaudhary PM;
XX
DR MPI: 1999-205191/17.
DR N-PSDB; AAX23425.
XX
PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
PS Claim 40; Fig 13B; 156pp; English.
XX
XX This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRI and TNRI3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/ active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the changer in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
CC
SQ Sequence 211 AA:

Query Match 93.7%; Score 1089; DB 20; Length 211;
Best Local Similarity 99.5%; Pred. No. 4.9e-105;
Matches 210; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 15 LVVSLGSMATLSAOPESOEELTAEDRRPEELNPQTESQDVVPLEQLVRRRAPKG 74
DB 1 LVVSLGSMATLSAOPESOEELTAEDRRPEELNPQTESQDVVPLEQLVRRRAPKG 60
OY 75 RRARRRAIAAHYEVHPRRGODGAGVGTVSGWEETKINSSSPRYRROI GEFVIRA 134
DB 61 RRARRRAIAAHYEVHPRRGODGAGVGTVSGWEETKINSSSPRYRROI GEFVIRA 120
OY 135 GLYYLTCQVHFDEGKAVYIKLIDLNVGVIALRCLLEFSATASSPGPOLRICOVSGLLPL 194
DB 121 GLYYLTCQVHFDEGKAVYIKLIDLNVGVIALRCLLEFSATASSPGPOLRICOVSGLLAL 180
OY 195 RGGSSLRIRTLPMAHKKAAPFLTYGFLFOVH 225
DB 181 RGGSSLRIRTLPMAHKKAAPFLTYGFLFOVH 211

RESULT 4
AAM29745
ID AAM29745 standard; Protein: 249 AA.
XX
AC AAM29745;
XX
DT 27-OCT-1998 (first entry)
XX
DE TNF related endothelium proliferative agent protein.
XX
KM TNF; endothelium proliferative agent; TRRPA; wound healing; cancer;

KM tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
XX
XX Homo sapiens.
XX OS
XX WO9835061-A2.
XX PN
XX 13-AUG-1998.
XX PD
XX 12-FEB-1998; 98WO-US02859.
XX PF
XX 10-FEB-1998; 98US-0021706.
XX PR
XX 12-FEB-1997; 97US-0798692.
XX PR
XX (ABBO) ABBOTT LAB.
XX PA
XX
XX PI Wiley SR;
XX PT
XX MPI: 1998-447255/38.
XX DR
XX N-PSDB; AAV47613.
XX DR
XX
XX Detecting nucleic acid encoding TRRPA - useful for diagnosis and
XX treatment of autoimmune disease, tumours and inflammation
XX
XX Claim 16; Page 123-4; 142pp; English.
XX
XX The TNF-related endothelium proliferative agent (TRRPA), or its
XX activators or agonists, are used to treat a deficit of TRRPA, e.g. to
XX promote wound healing or tissue grafting, by promoting vascularisation,
XX also to induce apoptosis for treating cancer and eliminating autoreactive
XX T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
XX TRRPA peptides can also be used to target cytotoxic agents or for
XX affinity isolation of the corresponding receptor, the nucleic acid for
XX which can be used to transform tumour cells to render them more
XX responsive to TRRPA and to screen for TRRPA mimics.
XX Ribozymes, antisense RNA, antibodies or peptides, are used to treat
XX TRRPA-associated diseases, e.g. tumours and metastases (by inhibiting
XX vascularisation), inflammation or a wide range of autoimmune conditions,
XX conditions involving abnormal stimulation of epithelial cells (e.g.
XX atherosclerosis), for birth control (inhibiting ovulation and placental
XX formation) or other angiogenic conditions (e.g. ulcers).
XX
SQ Sequence 249 AA:

Query Match 87.8%; Score 1020; DB 19; Length 249;
Best Local Similarity 88.8%; Pred. No. 9.5e-98;
Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

OY 2 LSLGIALAGLLLVVSLGSMATLSAOPESOEELTAEDRRPEELNPQTESQDVVPL 61
DB 26 LGLGIALAGLLLVVSLGSMATLSAOPESOEELTAEDRRPEELNPQTESQDVVPL 85
OY 62 EQLVRRRAPKGRKRRARRAIAAHYEVHPRRGODGAGVGTVSGWEETKINSSSP 121
DB 86 NRIIVRRRAPKGRKRRARRAIAAHYEVHPRRGODGAGVGTVSGWEETKINSSSP 145
OY 122 YDROI GEFVIRAGLYLTCQVHFDEGKAVYIKLIDLNVGVIALRCLLEFSATASSPGP 181
DB 146 YNRROI GEFVIRAGLYLTCQVHFDEGKAVYIKLIDLNVGVIALRCLLEFSATASSPGP 205
OY 182 QLRICOVSGLLLRGGSSLRIRTLPMAHKKAAPFLTYGFLFOVH 225
DB 206 QLRICOVSGLLLRGGSSLRIRTLPMAHKKAAPFLTYGFLFOVH 249

RESULT 5
AAV09369
ID AAV09369 standard; Protein: 249 AA.
XX
AC AAV09369;
XX
DT 15-JUL-1999 (first entry)
XX
XX Human tumour necrosis factor Apo-3 ligand protein sequence.

XX Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
KM NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
KM cancer.
XX
OS Homo sapiens.
XX
PN MO9919490-A1.
XX
PD 22-APR-1999.
XX
PF 09-OCT-1998; 98WO-US21407.
XX
PR 17-DEC-1997; 97US-0069862.
PR 10-OCT-1997; 97US-0062037.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Marsters SA, Pictl R;
XX
DR WPI; 1999-287982/24.
DR N-PSDB; AAX56000.
XX
PT New human Apo-3 ligand (a tumour necrosis factor) homologue
XX
PS Claim 1; Fig 1; 74pp; English.
XX
CC The present sequence represents a human tumour necrosis factor (TNF)
CC and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
CC cytostatic activity. Apo-3 ligand can be used to induce apoptosis in
CC mammalian cancer cells, to induce NF-kappaB-dependent transcription and
CC to induce JNK/SAPK-dependent responses in mammalian cells.
XX
SQ Sequence 249 AA:

Query Match 87.8%; Score 1020; DB 20; Length 249;
Best Local Similarity 88.8%; Pred. No. 9.5e-98;
Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 2 LSLGLALACGLLLVNVSLGSMATLSAQPSEELTAEDRREPELNPQTEESQDVVFL 61
DB 26 LGGLALACGLLLVNVSLGSRASLSAQEPQAEELVAEEDDPSELNPQTEESQDPAFL 85
QY 62 EQLVPRRSAPKGRKARRRAIAHVEVHPRPGDGAQAGVDGTVSGWEETKINSSPLR 121
DB 86 NRLVPRRSAPKGRKARRRAIAHVEVHPRPGDGAQAGVDGTVSGWEETKINSSPLR 145
QY 122 YDRQIGFTYRAGLYLYLCVHDEGKAVYLRKLDLVNGLALRCLFEFSATTAASSGP 181
DB 146 YNRQIGFTYRAGLYLYLCVHDEGKAVYLRKLDLVNGLALRCLFEFSATTAASSGP 205
QY 182 QLRLQVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYGFLQVH 225
DB 206 QLRLQVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYGFLQVH 249

RESULT 6
AAB07526
ID AAB07526 standard; protein; 249 AA.
XX
AC AAB07526;
XX
DT 20-OCT-2000 (first entry)
XX
DE Amino acid sequence of a soluble recombinant human TWEAK protein.
XX
KM TWEAK protein; immunological disorder; immune response; inflammation;
KM TWEAK blocking agent; autoimmune disease; organ transplant rejection;
KM Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour.
XX
OS Homo sapiens.
XX
PN WO200042073-A1.

XX 20-JUL-2000.
PD 14-JAN-2000; 2000WO-US01044.
XX
PF 15-JAN-1999; 99US-0116168.
XX
PR (BIOJ) BIOGEN INC.
XX
PA Renner P;
XX
PI WPI; 2000-476036/41.
XX
DR
XX
PT Preventing and treating immune responses using modulators, especially
PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for
PT treating e.g. inflammation and graft versus host disease -
XX
PS Disclosure; Fig 1; 45pp; English.
XX
XX
CC The present sequence represents a TWEAK protein. The specification
CC describes a method for preventing or treating an immunological
CC disorder and/or inhibiting an immune response in an animal. The
CC method comprises administering a TWEAK blocking agent. The method may
CC be used for preventing and treating immune disorders associated with
CC inappropriate expression and/or activity of TWEAK. These disorders
CC include autoimmune diseases, acute and chronic inflammation, organ
CC transplant rejection, Graft-versus-host disease (GVHD), lymphoid cell
CC malignancies (as seen in human immunodeficiency virus (HIV)
CC infections) and failure of the immune response to tumour growth.
XX
SQ Sequence 249 AA:

Query Match 87.8%; Score 1020; DB 21; Length 249;
Best Local Similarity 88.8%; Pred. No. 9.5e-98;
Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 2 LSLGLALACGLLLVNVSLGSMATLSAQPSEELTAEDRREPELNPQTEESQDVVFL 61
DB 26 LGGLALACGLLLVNVSLGSRASLSAQEPQAEELVAEEDDPSELNPQTEESQDPAFL 85
QY 62 EQLVPRRSAPKGRKARRRAIAHVEVHPRPGDGAQAGVDGTVSGWEETKINSSPLR 121
DB 86 NRLVPRRSAPKGRKARRRAIAHVEVHPRPGDGAQAGVDGTVSGWEETKINSSPLR 145
QY 122 YDRQIGFTYRAGLYLYLCVHDEGKAVYLRKLDLVNGLALRCLFEFSATTAASSGP 181
DB 146 YNRQIGFTYRAGLYLYLCVHDEGKAVYLRKLDLVNGLALRCLFEFSATTAASSGP 205
QY 182 QLRLQVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYGFLQVH 225
DB 206 QLRLQVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYGFLQVH 249

RESULT 7
AAV95338
ID AAV95338 standard; protein; 249 AA.
XX
AC AAV95338;
XX
DT 25-SEP-2000 (first entry)
XX
DE Human PRO207 antitumour protein.
XX
KM PRO207; human; antitumour; tumour; therapy; cytostatic;
KM breast cancer; ovarian cancer; renal cancer; colorectal cancer;
KM uterine cancer; prostate cancer; lung cancer; bladder cancer;
KM central nervous system cancer; melanoma; leukaemia; neoplasm.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..40

FT	Protein	/label= Signal_peptide
FT		41..249
FT		/label= PRO207
FT	Modified-site	27..33
FT		/note= "N-myristoylation"
FT	Modified-site	29..35
FT		/note= "N-myristoylation"
FT	Modified-site	36..42
FT		/note= "N-myristoylation"
FT	Modified-site	45..51
FT		/note= "N-myristoylation"
FT	Modified-site	118..124
FT		/note= "N-myristoylation"
FT	Modified-site	121..127
FT		/note= "N-myristoylation"
FT	Modified-site	125..131
FT		/note= "N-myristoylation"
FT	Modified-site	128..134
FT		/note= "N-myristoylation"
FT	Modified-site	139..143
FT		/note= "asn is N-glycosylated"
FT	Modified-site	10..14
FT		/note= "amidation"
FT	Modified-site	97..101
FT		/note= "amidation"
FT	Peptide	24..35
FT		/note= "prokaryotic membrane lipoprotein lipid"
XX		
XX	WO200037638-A2.	
XX	29-JUN-2000.	
XX		
XX	02-DEC-1999;	99MO-US25865.
XX		
XX	22-DEC-1998;	98US-0113296.
XX	08-MAR-1999;	99MO-US05028.
XX	21-APR-1999;	99US-0130232.
XX	28-APR-1999;	99US-0131445.
XX	14-MAY-1999;	99US-0134287.
XX	20-JUL-1999;	99US-0144758.
XX	26-JUL-1999;	99US-0145698.
XX	15-SEP-1999;	99MO-US21090.
XX	15-SEP-1999;	99MO-US21547.
XX		
XX	(GEPH) GENENTECH INC.	
XX	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA;	
XX	Napier MA, Pitti RM, Wood WI;	
XX	WPI; 2000-442668/38.	
XX	N-PSDB; AAA49717.	
XX		
XX	Novel composition to inhibit neoplastic cell growth or for treating	
XX	tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,	
XX	PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or	
XX	PRO866	
XX	Claim 19; Fig 4; 172pp; English.	
XX		
XX	The present sequence is that of human antitumour protein PRO207,	
XX	as deduced from a foetal kidney cDNA clone (see AAA49717). PRO207	
XX	shows amino acid sequence identity to tumour necrosis factor family	
XX	members, especially human lymphotxin-beta (23.4%) and human CD40	
XX	ligand (19.8%). Mol.wt. is 27,216. A claimed method for inhibiting	
XX	the growth of a tumour cell comprises exposing the tumor cell	
XX	to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301,	
XX	PRO356, PRO362, PRO356, PRO509 or PRO866 (see AAY9537-49), their	
XX	agonists or chimeric polypeptides incorporating them. The tumour	
XX	is especially a cancer selected from breast, ovarian, renal,	
XX	colorectal, uterine, prostate, lung, bladder and central nervous	
XX	system cancer, melanoma and leukaemia. Methods for the recombinant	
XX	expression of the antitumour proteins are also provided.	

Sequence	249 AA;	87.8%; Score 1020;	DB 21;	Length 249;
Query Match	88.8%;	Pred. No. 9.5e-98;		
Best Local Similarity	88.8%;	Pred. No. 9.5e-98;		
Matches 199;	Conservative 9;	Mismatches 16;	Indels 0;	Gaps
OY	2	LSTGLATACLGILLVYVSLGSMATISAEPSQDELTAADREPELNPOTFEESODVYPL	61	
DB	26	LGLGLATACLGILLVYVSLGSRASISAEPSQDELTAADREPELNPOTFEESODVYPL	85	
OY	62	EOLVPRRSAPKGRKAPRRRAIAAHYEVHPPRGDDGAGVDTGVSQMEETKINSSPLR	121	
DB	86	NRLVPRRSAPKGRKTRARRAIAAHYEVHPPRGDDGAGVDTGVSQMEETKINSSPLR	145	
OY	122	YDROIGETFTVIRAGLYLYLCOVHDEGKAVYIKLDELVNGVALRCLSEFSATTAASSGP	181	
DB	146	YNROIGETFTVIRAGLYLYLCOVHDEGKAVYIKLDELVNGVALRCLSEFSATTAASSGP	205	
OY	182	QIRLCQVSGCLPLRPGSSLRIRTPMALHKAAPLTYGGLQVH	225	
DB	206	QIRLCQVSGLLALRPGSSLRIRTPMALHKAAPLTYGGLQVH	249	
RESULT 8				
AAE00891				
ID	AAE00891	standard; Protein: 249 AA.		
XX	AAE00891;			
AC				
XX				
DT	04-JUL-2001	(first entry)		
XX				
DE	Human TREPA (TNF related endothelium proliferative agent).			
XX				
KM	Human; tumour necrosis factor; TNF; angiogenesis; wound healing;			
KW	TREPA; TNF related endothelium proliferative agent; tumour; metastasis;			
KM	grafting; vulnerary.			
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	Domain	98..249		
FT		/label= Extracellular_domain		
XX				
PN	US6207642-B1.			
XX				
PD	27-MAR-2001.			
XX				
PF	26-JUN-1998;	98US-0105343.		
XX				
PR	12-FEB-1997;	97US-0798692.		
PR	10-FEB-1998;	98US-0021706.		
XX				
PA	(ABBO) ABBOTT LAB.			
XX				
PI	Wiley SR;			
XX				
DR	WPI: 2001-280760/29.			
XX				
DR	N-PSDB: AAD04350.			
XX				
PT	Inducing angiogenesis in mammal at desired sites for promoting wound			
PT	healing, by administering soluble fragment of extracellular domain of			
PT	tumour necrosis factor related endothelium proliferative agent protein			
XX				
XX				
PS	Claim 1; Column 75-76; 53pp; English.			
XX				
CC	The present invention relates to extracellular signal molecules,			
CC	particularly members of tumour necrosis factor (TNF) family molecules			
CC	designated as TREPA (TNF related endothelium proliferative agent).			
CC	Soluble biologically active TREPA are used to treat TREPA-associated			
CC	diseases, tumours or metastases. TREPA is used for inducing angiogenesis			
CC	in human for promoting wound healing and for vascularising grafted			
CC	tissue for successful grafting and to promote tissue grafts.			

CC The present amino acid sequence is clone ID #690050 human TREPA.
 XX
 SQ Sequence 249 AA;
 Query Match 87.8%; Score 1020; DB 23; Length 249;
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;
 XX
 QY 2 LSLGLALACGLLLVNVSLGSMATLSAQEPSOEELTAEDRRPEELNPQTEESQDVVPL 61
 DB 26 LGGLALACGLLLVNVSLGSRASLSAQEPQELVAEEDQDSELPQTEESQDPAFL 85
 QY 62 EQLVPRRSAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDGVSQWEETKINSSPLR 121
 DB 86 NRLVPRRSAPKGRKTRARRAIAAHYEVHPRPGDGAQAGVDGVSQWEETKINSSPLR 145
 QY 122 YDRQIGFTYIRAGLYLYLCQVHDEGKAVYTKLDLVNGLALRCLEFSATASSPCP 181
 DB 146 YNRQIGFTYIRAGLYLYLCQVHDEGKAVYTKLDLVNGLALRCLEFSATASSPCP 205
 QY 182 QLRICQVSGLLPRGSSSLRIRLPMNAHLKAAPFLTYFGLFOVH 225
 DB 206 QLRICQVSGLLALRPGSSSLRIRLPMNAHLKAAPFLTYFGLFOVH 249
 RESULT 9
 AAU86129
 ID AAU86129 standard; Protein; 249 AA.
 XX
 AC AAU86129;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE Human PRO207 polypeptide.
 XX
 XX Human: PRO: benign tumour; malignant tumour; lymphoid malignancy;
 KM Leukemia; neuronal disorder; stromal disorder; blastocoelec disorder;
 KM Inflammatory disorder; Immune disorder; angioleptic disorder;
 KM Cytostatic; neuroprotective.
 XX
 OS Homo sapiens.
 XX
 PN WO200153486-A1.
 XX
 PD 26-JUL-2001.
 XX
 PF 11-FEB-2000; 2000WO-US03565.
 XX
 PR 08-MAR-1999; 99WO-US05028.
 PR 11-MAR-1999; 99US-123972P.
 PR 11-MAY-1999; 99US-133459P.
 PR 02-JUN-1999; 99WO-US12252.
 PR 22-JUN-1999; 99US-140650P.
 PR 22-JUN-1999; 99US-140653P.
 PR 20-JUL-1999; 99US-144758P.
 PR 26-JUL-1999; 99US-145688P.
 PR 28-JUL-1999; 99US-146222P.
 PR 17-AUG-1999; 99US-149395P.
 PR 31-AUG-1999; 99US-151689P.
 PR 01-SEP-1999; 99WO-US20111.
 PR 15-SEP-1999; 99WO-US21090.
 PR 30-NOV-1999; 99WO-US28313.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 05-JAN-2000; 2000WO-US00219.
 XX
 XX (GETH) GENENTECH INC.
 XX
 XX Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
 PI Marsters SA, Pan J, Pittl RM, Roy MA, Smith V, Stone DM;
 PI Watanabe CK, Wood WI;
 XX
 DR WPI; 2002-205567/26.

DR N-PSDB; ABK40255.
 XX
 PT Thirty five nucleic acids encoding PRO polypeptides, useful for
 PT treating benign or malignant tumours, leukaemias and lymphoid
 PT malignancies, inflammatory, angiogenic and immunologic disorders -
 XX
 PS Claim 61; Fig 4; 302pp; English.
 XX
 CC The present invention relates to the isolation of novel human PRO
 CC polypeptides and the polynucleotide sequences encoding them. The
 CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
 CC useful for treating benign or malignant tumours (e.g. renal, kidney,
 CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
 CC disorders such as neuronal, glial, astrocytic, hypochalamic, glandular,
 CC macrophagal, stromal and blastocoelec disorders, inflammatory, immune
 CC and angiogenic disorders. The polynucleotide sequences are also
 CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO
 CC polypeptides of the invention.
 CC
 SQ Sequence 249 AA;
 XX
 Query Match 87.8%; Score 1020; DB 23; Length 249;
 Best Local Similarity 88.8%; Pred. No. 9.5e-98;
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;
 XX
 QY 2 LSLGLALACGLLLVNVSLGSMATLSAQEPSOEELTAEDRRPEELNPQTEESQDVVPL 61
 DB 26 LGGLALACGLLLVNVSLGSRASLSAQEPQELVAEEDQDSELPQTEESQDPAFL 85
 QY 62 EQLVPRRSAPKGRKARPRRAIAAHYEVHPRPGDGAQAGVDGVSQWEETKINSSPLR 121
 DB 86 NRLVPRRSAPKGRKTRARRAIAAHYEVHPRPGDGAQAGVDGVSQWEETKINSSPLR 145
 QY 122 YDRQIGFTYIRAGLYLYLCQVHDEGKAVYTKLDLVNGLALRCLEFSATASSPCP 181
 DB 146 YNRQIGFTYIRAGLYLYLCQVHDEGKAVYTKLDLVNGLALRCLEFSATASSPCP 205
 QY 182 QLRICQVSGLLPRGSSSLRIRLPMNAHLKAAPFLTYFGLFOVH 225
 DB 206 QLRICQVSGLLALRPGSSSLRIRLPMNAHLKAAPFLTYFGLFOVH 249
 RESULT 10
 AAM47525
 ID AAM47525 standard; Protein; 284 AA.
 XX
 AC AAM47525;
 XX
 DT 21-JUL-1998 (first entry)
 XX
 DE Homo sapiens tumour necrosis factor related ligand (TRELL).
 XX
 XX TRELL: tumour necrosis factor related ligand; tnf; treatment;
 KM cancer; autoimmune disease; immune system; stimulation; suppression;
 KM graft rejection.
 XX
 OS Homo sapiens.
 XX
 PN WO9805783-A1.
 XX
 PD 12-FEB-1998.
 XX
 PF 07-AUG-1997; 97WO-US13945.
 XX
 PR 18-MAR-1997; 97US-0040820.
 PR 07-AUG-1996; 96US-0023541.
 PR 18-OCT-1996; 96US-0028515.
 XX
 XX (BIOI) BIOGEN INC.
 PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.
 XX
 PI Browning JL, Chicheportliche Y;

DR WP1: 1998-145619/13.
 DR N-PSDB: AAV18600.
 PT Tumour necrosis factor related ligand - useful for, e.g. treating
 cancer, auto-immune disease and immune responses to tissue grafts
 XX
 PS Claim 12: Pages 50-51; 69pp; English.
 CC The sequence is that of human tumour necrosis factor related
 CC ligand (TRELL). TRELL or active fragments can be included with a
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune
 CC diseases or immune responses to tissue grafts, or to stimulate or
 CC suppress the immune system. It is useful to screen for TRELL
 CC receptors, by labelling with a detectable label and screening
 CC compositions for binding. Agents interfering with TRELL-receptor
 CC binding can also be screened for, can then be administered,
 CC optionally with interferon- gamma, to induce cell death or
 CC treat, suppress or alter immune responses (especially involving human
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its
 CC receptor. It's coding sequence can be used in gene therapy for
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
 CC autoimmune and inflammatory diseases or inherited genetic disorders,
 CC by introducing into cells, and expressing, therapeutically effective
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
 CC It may also be of use in the preparation of prepare probes for
 CC screening natural/synthetic DNAs for TRELL-encoding sequences
 CC and for antisense therapy.
 CC
 XX Sequence 284 AA:
 S0
 Query Match 87.8%; Score 1020; DB 19; Length 284;
 Best Local Similarity 88.8%; Pred. No. 1.1e-97;
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;

QY 2 LSLGLALACILLLVVSLSGSMATLSAQEPSELTAEEDRREPPLNPQTEESQDVVPL 61
 DB 61 LGLLALACILLLVVSLSGSMATLSAQEPSELTAEEDRREPPLNPQTEESQDVVPL 120
 QY 62 EQLVPRRSAPKGRKARRRAIAHYEVHPRPGDGAQAGVDTVSGMEARINSSPLR 121
 DB 121 NRLVPRRSAPKGRKARRRAIAHYEVHPRPGDGAQAGVDTVSGMEARINSSPLR 180
 QY 122 YDRQIGFEVTIRAGLYLYLCQVHDEGKAVYLKIDLVLNGVLAIRCLSEFSATASPGP 181
 DB 181 YNRQIGFEVTIRAGLYLYLCQVHDEGKAVYLKIDLVLNGVLAIRCLSEFSATASPGP 240
 QY 182 QLRICQVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYFGLFQVH 225
 DB 241 QLRICQVSGLLPLRPGSSLRIRTLPMWHLKAAPFLTYFGLFQVH 284

RESULT 11
 AAW93590
 ID AAW93590 standard; Protein; 208 AA.
 AC AAW93590;
 XX
 XX 18-JUN-1999 (first entry)
 DE Human TNRL3 protein.
 XX
 KW Tumour necrosis factor receptor; signal transducer molecule; TNF; ApO4;
 KW developmental abnormality; gestational abnormality; prostate cancer;
 KW ApO6; ApO8; ApO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KW apoptosis; human.
 XX
 XX Homo sapiens.
 OS
 XX
 XX W09911791-A2.
 PN
 XX 11-MAR-1999.
 PD
 XX

PF 04-SEP-1998; 98W0-US18393.
 XX
 PR 05-SEP-1997; 97US-0924634.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 XX Chaudhary PM;
 XX
 DR WP1: 1999-205191/17.
 DR N-PSDB: AAX23424.
 XX
 PT New Tumour Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 XX
 PS Claim 40; Fig 13A; 156pp; English.
 CC This invention describes isolated Tumour Necrosis Factor (TNF) family
 CC receptor polypeptides: ApO4, ApO6, ApO8 and ApO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. ApO4 is useful for diagnosing prostate cancer
 CC by determining levels of ApO4 in an individual. Prostate cancer can also
 CC be treated using ApO4 selective binding agents linked to a therapeutic
 CC moiety. ApO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. ApO4 polypeptides/active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the changer in ApO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using ApO4 polypeptides/active
 CC fragments and ApO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of ApO4 and detecting a change in level of ApO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. ApO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. ApO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.
 CC
 XX Sequence 208 AA:
 S0
 Query Match 81.8%; Score 951; DB 20; Length 208;
 Best Local Similarity 88.9%; Pred. No. 1.1e-90;
 Matches 185; Conservative 8; Mismatches 15; Indels 0; Gaps 0;

QY 18 VSLGSMATLSAQEPSELTAEEDRREPPLNPQTEESQDVVPLFQLVPRRSAPKGRKA 77
 DB 1 VSLGSMATLSAQEPSELTAEEDRREPPLNPQTEESQDVVPLFQLVPRRSAPKGRKT 60
 QY 78 RPRRAIAHYEVHPRPGDGAQAGVDTVSGMEARINSSPLRDRQIGFEVTIRAGLY 137
 DB 61 RARRRAIAHYEVHPRPGDGAQAGVDTVSGMEARINSSPLRIRYRQIGFEVTIRAGLY 120
 QY 138 YLYCQVHDEGKAVYLKIDLVLNGVLAIRCLSEFSATASPGQRLRCQVSGLLPLRPG 197
 DB 121 YLYCQVHDEGKAVYLKIDLVLNGVLAIRCLSEFSATASPGQRLRCQVSGLLPLRPG 180
 QY 198 SSLRIRTLPMWHLKAAPFLTYFGLFQVH 225
 DB 181 SSLRIRTLPMWHLKAAPFLTYFGLFQVH 208

RESULT 12
 AAW03499
 ID AAW03499 standard; Protein; 273 AA.
 AC AAW03499;
 XX
 XX 26-SEP-2001 (first entry)
 DE TWEAK extracellular domain-containing fusion protein.
 XX
 XX TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;
 KW

KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
 KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
 KM ruberosis; uveitis; macular degeneration; arthritis; rheumatism;
 KM corneal graft neovascularisation; psoriasis; metastatic condition;
 KM malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
 KM preneoplastic condition; myocardial angiogenesis; wound granulation;
 KM scleroderma; vascular adhesion; telangiectasia; ischaemia; human;
 KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;
 KM peripheral atherosclerosis; PDC409-L2-TWEAK; TWEAK receptor; TWEAKR;
 fusion protein.
 XX
 OS Homo sapiens.
 OS Synthetic.
 PN WO200145730-A2.
 XX
 XX 28-JUN-2001.
 PD
 XX 19-DEC-2000; 2000WO-US34755.
 PF
 XX 20-DEC-1999; 99US-0172878.
 PR 10-MAY-2000; 2000US-0203347.
 XX
 XX (IMMUNEX) IMMUNEX CORP.
 PA
 XX Willey SR.
 PI
 XX N-PDB: AAS03964.
 DR
 XX
 XX
 PS Example 1; Page 41; 46pp; English.
 CC The sequence represents a fusion protein encoded by the
 CC expression vector PDC409-L2-TWEAK. The fusion protein comprises a
 CC growth hormone leader, a leucine zipper multimerisation domain, and
 CC the extracellular domain of human TWEAK. The fusion protein was
 CC used in the isolation of human TWEAK receptor (TWEAKR)-expressing
 CC clones from a COS cell human cDNA library. The TWEAK protein is
 CC a member of the tumour necrosis factor (TNF) family and induces
 CC angiogenesis. TWEAKR may therefore be used to screen for and
 CC develop TWEAKR agonists and antagonists for the modulation of
 CC angiogenesis, to be used in the treatment and diagnosis of human disease.
 CC The disorders mediated by angiogenesis include ocular disorders
 CC characterised by ocular neovascularisation such as diabetic retinopathy,
 CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,
 CC retrolental fibroplasia, ruberosis, uveitis, macular degeneration and
 CC corneal graft neovascularisation, and inflammatory diseases such as
 CC arthritis, rheumatism and psoriasis. Other treatable diseases include
 CC malignant and metastatic conditions such as sarcomas and carcinomas,
 CC benign tumours and preneoplastic conditions, myocardial angiogenesis,
 CC haemophilic joints, scleroderma, vascular adhesions, atherosclerotic
 CC plaque neovascularisation, telangiectasia, wound granulation, coronary
 CC atherosclerosis, peripheral atherosclerosis and ischaemia.
 CC
 XX Sequence 273 AA:
 SO
 Query Match 81.8%; Score 951; DB 22; Length 273;
 Best Local Similarity 88.9%; Pred. No. 1.6e-90;
 Matches 184; Conservative 9; Mismatches 14; Indels 0; Gaps 0;

QY 139 LYCVHFEDEKAVYKLDLVNGVLAARCLSEFSATASSPGQRLRCQVSGLLRPGS 198
 DB 187 LYCVHFEDEKAVYKLDLVNGVLAARCLSEFSATASSPGQRLRCQVSGLLRPGS 246
 QY 199 SLRIRLPMANHLKAAPFLTYFGLFQVH 225
 DB 247 SLRIRLPMANHLKAAPFLTYFGLFQVH 273
 RESULT 13
 AAW29746
 ID AAW29746 standard; Protein: 189 AA.
 XX
 AC AAW29746;
 XX
 DT 27-OCT-1998 (first entry)
 XX
 DE TNF related endothelium proliferative agent protein 2.
 XX
 KM TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
 KM tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
 XX
 OS Homo sapiens.
 XX
 XX WO9835061-A2.
 PN
 XX 13-AUG-1998.
 PD
 XX 12-FEB-1998; 98WO-US02859.
 PF
 XX 10-FEB-1998; 98US-0021706.
 PR 12-FEB-1997; 97US-0798692.
 XX
 XX (ABBO) ABBOTT LAB.
 PA
 XX
 XX Willey SR.
 PI
 XX WPI: 1998-447255/38.
 DR
 XX
 XX
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
 treatment of autoimmune disease, tumours and inflammation
 XX
 PS Claim 16; Page 125-6; 142pp; English.
 XX
 XX The TNF-related endothelium proliferative agent (TREPA), or its
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
 CC promote wound healing or tissue grafting, by promoting vascularisation,
 CC also to induce apoptosis for treating cancer and eliminating autoreactive
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
 CC TREPA peptides can also be used to target cytotoxic agents or for
 CC affinity isolation of the corresponding receptor, the nucleic acid for
 CC which can be used to transform tumour cells to render them more
 CC responsive to TREPA and to screen for TREPA mimics.
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
 CC vascularisation), inflammation or a wide range of autoimmune conditions,
 CC conditions involving abnormal stimulation of epithelial cells (e.g.
 CC atherosclerosis), for birth control (inhibiting ovulation and placental
 CC formation) or other angiogenic conditions (e.g. ulcers).
 CC
 XX Sequence 189 AA:
 SO
 Query Match 63.9%; Score 742; DB 19; Length 189;
 Best Local Similarity 92.8%; Pred. No. 5.7e-69;
 Matches 142; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

THIS PAGE BLANK (USPTO)

[illegible]

```

FEATURES
  source
    location/Qualifiers
      1. .918
        /organism="Mus musculus"
        /strain="FVB/N"
        /db_xref="taxon:10090"
        /clone_image="IMAGE:4206595"
        /clone_lib="NCI CGAP Co24"
        /lab_host="DH10B (T1 phage-resistant)"
        /note="Organ: Colon; Vector: pCMV-SPORT6; Site:1: NotI;
        Site:2: SalI; Cloned unidirectionally. Primer: Oligo dT

```

Average insert size 1.6 kb. Constructed by Life Technologies. Note: this is a NCICGAP Library.			
BASE COUNT	153 a	292 c	268 g
ORIGIN			204 t
			I others
Query Match	53.6%	Score 625.8;	DB 12; Length 918;
Best Local Similarity	9.2%	Pred. No. 7.2e-152;	
Matches 660; Conservative	0; Mismatches 2;	Indels 3;	Gaps 3

OY	22	GGGGCTCGCCCTGGGACCTCCCTGCTGGTGGTCAAGCCTGGGGAGCTGGGACAGCTGTC	81
Db	1	GCTGGCTCCTCCTGGGCTCCTCCTGCTGGTGGTCAAGCCTGGGGAGCTGGGACAGCTGTC	60
OY	82	TGCCCCAGAGCCCTTCTCAGGAGGACCTGCACAGAGAGACCGCCGGGAGCCCTCGAACT	141
Db	61	TGCCCCAGAGCCCTTCTCAGGAGGACCTGCACAGAGAGACCGCCGGGAGCCCTCGAACT	120
OY	142	GAATCCCCAGACAGAGAGAAAGCCAGAGATGTTGTTACTTCTTGGAACAATCTAGTCGGCC	201
Db	121	GAATCCCCAGACAGAGAGAAAGCCAGAGATGTTGTTACTTCTTGGAACAATCTAGTCGGCC	180
OY	202	TGGAAGAGTGTCTCTTAAAGGCCGGAGAGCCGGGCTCGCCAGAGTATTGCAGCCATT	261
Db	181	TGGAAGAGTGTCTCTTAAAGGCCGGAGAGCCGGGCTCGCCAGAGTATTGCAGCCATT	239
OY	262	TGAGTTTATCTCTCGGCCAGACACAGATGAGACACAAGAGTGTGTGATGGACAGTGAG	321
Db	240	TGAGTTTATCTCTCGGCCAGACACAGATGAGACACAAGAGTGTGTGATGGAGACAGTGAG	299
OY	322	TGGCTGGGAGAGAGACCAAAATCAAGAGCTCCAGGCTCTGGCGCTACAGACGGCAGATTGG	381
Db	300	TGGCTGGGAGAGAGACCAAAATCAAGAGCTCCAGGCTCTGGCGCTACAGACGGCAGATTGG	359
OY	382	GSAAATTACAGTCAATCAGAGGCTGGGCTCTACTACTGTACTGTCAAGGTGCACCTTGATGA	441
Db	360	GSAAATTACAGTCAATCAGAGGCTGGGCTCTACTACTGTACTGTCAAGGTGCACCTTGATGA	419
OY	442	GGGAAGGCTGCTCTACCTGAACTGGAGCTGGTGGGAACGGGTGTGCTGGCCCTGGCGTG	501
Db	420	GGGAAGGCTGCTCTACCTGAACTGGAGCTGGTGGGAACGGGTGTGCTGGCCCTGGCGTG	479
OY	502	CCTGGAAGAAATCTCAGCCACAGCAGCAAGCTCTCCTGGGGCCCAAGCTCCGTTTGCCA	561
Db	480	CCTGGAAGAAATCTCAGCCACAGCAGCAAGCTCTCCTGGGGCCCAAGCTCCGTTTGCCA	539
OY	582	GGTGTCTGGGCGTGTGGCGCTGGGCGAGGGGCTTCCCTTGGATGCCAGCCTCCCTGG	621
Db	560	GGTGTCTGGGCGTGTGGCGCTGGGCGAGGGGCTTCCCTTGGATGCCAGCCTCCCTGG	598
OY	622	GGCCTATCTTAAAGCTGGCCCTTCTTAACCTACTTGTGGACTCTTTCAGAGTTACAGTAGG	681
Db	599	GGCCTATCTTAAAGCTGGCGTG-CCCTTCTTAACCTAATTTGGACTCTTTCAGAGTTACAGTAGG	657
OY	682	GGCCT 686	
Db	658	GGCCT 662	

RESULT		3
LOCUS	B1871711	
DEFINITION	B1871711 731 bp mRNA linear EST 11-OCT-2001 60339582SEF1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:5405478 5' , mRNA sequence.	
ACCESSION	B1871711	
VERSION	B1871711.1 GI:16045386	
KEYWORDS	EST.	
SOURCE	human. human.	
ORGANISM	Homo sapiens	
Eukaryota:	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia:	Eutheria; Primates; Catarrhini; Homnidae; Homo.	
I (bases 1 to 731)		
NIH-MGC http://mgc.ncl.nih.gov/. National Institutes of Health. Mammalian Gene Collection (MGC)		
REFERENCE		
AUTHORS		
TITLE		

Db 241 AGCCATTATGAATTCATCCACGACCTGGACAGCGAGCGAGAGGTGTGACGG 300
 QY 313 GACAGTAGAGTGGCTGGGAGAGACCAAAATCAACAGCTCCAGCTGCGCTACGACGG 372
 Db 301 GACAGTAGAGTGGCTGGGAGAGACCAAAATCAACAGCTCCAGCTGCGCTACGACGG 360
 QY 373 CCAGATTGGGAAATTTACAGCTCAGGCTGGGCTCTACTACTCTACTCTGTCAGGTGCA 432
 Db 361 CCAGATCGGGGAGTTTATAGTCACCGGGCTGGGCTCTACTACTACTCTGTCAGGTGCA 420
 QY 433 CTTTATGAGGAGAAAGGCTGTCTACCTGAGAGCTGAGCTTGTCTGTGAACGCTGTCTGCG 492
 Db 421 CTTTATGAGGAGAAAGGCTGTCTACCTGAGAGCTGAGCTTGTCTGTGAAGGTGTCTGCG 480
 QY 493 CCGGAGCTGGCTGGAAGATTCCTCAGCCACAGACAGCTCTCTGCGGCCCTGCTCG 552
 Db 481 CCGGAGCTGGCTGGAAGATTCCTCAGCCACAGCTCTCTGCGGCCCTGCTCGGCCCTGCG 540
 QY 553 TTTGTGACAGAGTGTCTGGGCTGTGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 612
 Db 541 CCGTGGCCAGAGTGTCTGGGCTGTGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 599
 QY 613 CCGTGGCCAGAGTGTCTGGGCTGTGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 672
 Db 600 CCGTGGCCAGAGTGTCTGGGCTGTGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 658
 QY 673 TCAGTGAAGGGGCTGTCTGGGCTGTGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 697
 Db 659 TCAGTGAAGGGGCTGTCTGGGCTGTGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 683

RESULT 5
 AM763237 561 bp mRNA linear EST 04-MAY-2000
 LOCUS ur70d09.y1 NCI-CGAP Mam3 Mus musculus cDNA clone IMAGE:315633 5'
 DEFINITION similar to TR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS
 ; mRNA sequence.
 AM763237
 AM763237.1 GI:7695174
 EST.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 561)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Other ESTs: ur70d09.x1
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
image.llnl.gov/image/html/iresources.shtml

MG1:1058389
 Seq primer: -40RP from Gibco
 High quality: sequence stop: 433.
 Location/Qualifiers
 1. 561

FEATURES
 SOURCE
 /organism="Mus musculus"
 /strain="129.C57BL/6J,FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:315633"
 /clone_lib="NCI-CGAP_Mam3"
 /tissue_type="tumor, gross tissue"
 /dev_stage="10 months"
 /lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-Sport6; Site.1: SalI;
 Site.2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigators
 providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
 Reference for Transgenic model: Xu et al., Nature Genetics
 22, 37-43 (1999)."
 BASE COUNT 108 a 158 c 194 g 100 t 1 others
 ORIGIN

Query Match 43.2%; Score 504; DB 10; Length 561;
 Best Local Similarity 99.8%; Pred. No. 3,1e-120;
 Matches 504; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGTGCTGAGCCCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 60
 Db 57 GGTGCTGAGCCCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 116
 QY 61 GGGAGAGCTGGGAGGAGCTGTCTGCCCAGAGAGCTTCTCAGAGAGCTGACAGCAGAGA 120
 Db 117 GGGAGAGCTGGGAGGAGCTGTCTGCCCAGAGAGCTTCTCAGAGAGCTGACAGCAGAGA 176
 QY 121 CCGCCGGGAGCCCTTGAATCTGAATCCACAGAGAGAGAGAGAGAGATGTGTACTTT 180
 Db 177 CCGCCGGGAGCCCTTGAATCTGAATCCACAGAGAGAGAGAGAGATGTGTACTTT 236
 QY 181 CTTGAGACACTAGTGTCCGCTCGAAGAGAGTCTCTTAAGGCGGAGAGGCGGCTGTG 240
 Db 237 CTTGAGACACTAGTGTCCGCTCGAAGAGAGTCTCTTAAGGCGGAGAGGCGGCTGTG 296
 QY 241 CCGAGCTATTGACAGCCATTATGAGGTTTATCCTCGGCGAGAGAGAGATGAGACAGC 300
 Db 297 CCGAGCTATTGACAGCCATTATGAGGTTTATCCTCGGCGAGAGAGAGATGAGACAGC 356
 QY 301 AGGTGTGATGGAGACAGTGTGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCT 360
 Db 357 AGGTGTGATGGAGACAGTGTGCTGGGAGAGACCAAAATCAACAGCTCCAGCCCTCT 416
 QY 361 GCGCTACGAGCCGCGAGTATGGGAAATTTACAGTCAAGAGGCTGGGCTCTACTACTGTGA 420
 Db 417 GCGCTACGAGCCGCGAGTATGGGAAATTTACAGTCAAGAGGCTGGGCTCTACTACTGTGA 476
 QY 421 CTGTGAGGTGCACTTTGATGAGGAGAGAGCTGTCTACCTGAAGCTGGACTGTGTGTA 480
 Db 477 CTGTGAGGTGCACTTTGATGAGGAGAGAGCTGTCTACCTGAAGCTGTGTGTA 536
 QY 481 CCGTGTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 505
 Db 537 CCGTGTGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 561

RESULT 6
 LOCUS BE628951 533 bp mRNA linear EST 25-AUG-2000
 DEFINITION uc30c03.y1 Soares_mammary_gland_MMLM3 Mus musculus cDNA clone
 IMAGE:337344 5' similar to TR:054907 054907 TNF-RELATED WEAK
 INDUCER OF APOPTOSIS ; mRNA sequence.
 BE628951
 BE628951.1 GI:9911639
 EST.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT

house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 533)
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 Tumor Gene Index
 Unpublished (1997)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (image.llnl.gov) for further information.
 MG1:1083048


```

|||||
44 GAACGAATCCAGACAGAAAGACAGATCTGCGCTTCTGAAACGACATGATT 103
197 CGGCTCGAAGAGTCTCTTAAGCCGAGAGCGGCTGCGGACTATTGACGC 256
104 CGGCTCGAAGAGTCTCTTAAGCCGAGAGCGGCTGCGGACTATTGACGC 163
257 CATATGAGTTCATCTCGGCGAGAGAGATGAGACACAGCAGTGTGATGGACA 316
164 CATATGAGTTCATCTCGGCGAGAGAGATGAGACACAGCAGTGTGATGGACA 223
317 GTGAGTGTGAGAGAGACCAAAATCAACAGCTCCAGCCCTTGGCTACGACGCCAG 376
224 TTGAGTGTGAGAGAGACCAAAATCAACAGCTCCAGCCCTTGGCTACGACGCCAG 283
377 ATTGGGGAATTTACATCATCAGAGGCTGGGCTTACTACTGTACTGTGATGACATT 436
284 ATCGGGAGTTATATATACCGCGGCTGGGCTTACTACTGTATCTGATGACATT 343
437 GATGAGGAAGAGCTGTCTACCTGAGCTGAGCTGTGATGAGAGTGTGATGAGCT 496
344 GATGAGGAAGAGCTGTCTACCTGAGCTGAGCTGTGATGAGAGTGTGATGAGCT 403
497 CGCTGCTGGAAGATTTCTACGACGACAGCAAGCTTCTGCGGCGGCTGCTTTG 556
404 CGCTGCTGGAAGATTTCTACGACGACGAGCTTCTGCGGCGGCTGCTTTG 463
557 TGGCAGTGTCTGGGCTGTGGGCTGGGCGGCGGCTTCTGCGGCGGCTGCTTTG 616
464 TGGCAGTGTCTGGGCTGTGGGCTGGGCGGCGGCTTCTGCGGCGGCTGCTTTG 523
617 CCTGAGCTCATCTTAAGGCTGCGGCTTCTGAGCTTGTGAGCTTGTGAGCTT 676
524 CCTGAGCTCATCTTAAGGCTGCGGCTTCTGAGCTTGTGAGCTTGTGAGCTT 583
677 TGAAGGCGCTTGTCTGCGGATTTCTTAACCTTCTGCGGCTGCGGAGCATCACACA 736
584 TGAAGGCGCTTGTCTGCGGATTTCTTAACCTTCTGCGGCTGCGGAGCATCACACA 636
737 CCTGCT---ACCCACCCCACTCTCTCCAGCTTCTGCGGCTGCGGAGCATCACACA 793
637 GCTGCTGAGGACCGGCTGCTCTGCGGCTGCGGAGCATCACACA 696
794 CTCTCC 799
697 CCTCTC 702

RESULT 14
AA221610 471 bp mRNA linear EST 13-FEB-1997
LOCUS my18d09.r1 Barstead mouse heart MRLB3 Mus musculus cDNA clone
DEFINITION IMAGE:696209 5', mRNA sequence.
ACCESSION AA221610
VERSION AA221610.1 GI:1840863
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 471)
AUTHORS Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
Theisinger,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
Waterston,R.
TITLE The WashU-HMI Mouse EST Project
JOURNAL Unpublished (1996)
COMMENT Contact: Maira M/Mouse EST Project
WashU-HMI Mouse EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800

```

```

Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (Info@image.lnl.gov) for further information.
MGI:429769
Seq primer: -28m13 rev2 ET from Amersham
High quality sequence stop: 301.
Location/Qualifiers
1..471
/organism="Mus musculus"
/strain="BALB/c"
/db_xref="taxon:10090"
/clone="IMAGE:696209"
/clone_lib="Barstead mouse heart MRLB3"
/sex="mixed"
/tissue_type="heart"
/dev_stage="6 weeks"
/lab_host="DH10B"
/note="Organ: heart; Vector: p773D-Pac (Pharmacia) with a
modified polylinker; Site.1: EcoRI; Site.2: NotI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACGATCTGAGAGTGGAGCGGCGGCTTTTCTTTTCTTTTCTTTTCTTTT
3']; double-stranded cDNA was ligated to Eco RI adaptors
[CTTGATTCGGTACC], digested with Not I and cloned into
the Not I and Eco RI sites of the modified p773 vector.
Library constructed by Bob Barstead."
BASE COUNT 91 a 141 c 128 g 111 t
ORIGIN
Query Match 37.5%; Score 437.8; DB 9; Length 471;
Best Local Similarity 97.0%; Pred. No. 5.1e-103;
Matches 457; Conservative 0; Mismatches 12; Indels 2; Gaps 1;
QY 274 TCGGCGACGACGATGAGACACAGAGTGTGATGGACGATGAGTGTGAGAGA 333
DB 3 TCGGATCCAAAGAGATGAGACACAGAGTGTGATGGACGATGAGTGTGAGAGA 62
QY 334 GACCAAAATCAAGCTCCAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 393
DB 63 GACCAAAATCAAGCTCCAGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 122
QY 394 CATCAGGCGTGGGCTTACTACTGATCTGATGATGATGATGATGATGATGATGAT 453
DB 123 CATCAGGCGTGGGCTTACTACTGATCTGATGATGATGATGATGATGATGATGAT 182
QY 454 CTACCTGAAGCTGAGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 513
DB 183 CTACCTGAAGCTGAGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 242
QY 514 CTGAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 573
DB 243 CTGAGCAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 302
QY 574 GTTCCGCTGCGGCGAGGCTTCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 633
DB 303 GTTGC--GCTGGCGGAGGCTTCTGCGGCTGCGGCTGCGGCTGCGGCTTCTTA 360
QY 634 GGTGCGGCTTCTTAAGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 693
DB 361 GGTGCGGCTTCTTAAGCTTGTGATGATGATGATGATGATGATGATGATGATGAT 420
QY 694 CCGATTTCTTAAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 744
DB 421 CCGATTTCTTAAGCTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 471

RESULT 15
BO671259 963 bp mRNA linear EST 15-JUL-2002
LOCUS BO671259
DEFINITION AGENCOURT.8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
ACCESSION BO671259
VERSION BO671259.1 GI:21782093

```

				KEYWORDS	ESt.
SOURCE	Homo sapiens			human.	
ORGANISM	Mammalia; Eutheria;			Mammalia; Chordata; Craniata; Vertebrata; Euteleostomi;	
REFERENCE	NIH-MGC http://mgi.nci.nih.gov/			(bases 1 to 963)	
AUTHORS	National Institutes of Health,			Mammalian Gene Collection (MGC)	
TITLE	Unpublished (1999)			Contact: Robert Strusberg, Ph.D.	
JOURNAL	Email: cga@remail.nih.gov			Tissue Procurement: ATCC	
COMMENT	CNA Library Preparation: Rubin Laboratory CNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov plate: LITCM2456 row: 1 column: 13 High quality sequence stop: S55. Location/Qualifiers				
FEATURES	source				
	1..963			/organism="Homo sapiens"	
	/db_xref="taxon:9606"			/clone="IMAGE:6274716"	
	/clone_1lb="NIH_MGC_102"			/tissue_type="epidermoid carcinoma, cell line"	
	/lab_host="DH10B (phage-resistant)"			"note":"Organ: salivary gland; Vector: pORF7; Site_1: XhoI, directionally cloned from EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(C). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)". Note: This is a NIH_MGC library."	
	Site_2: EcoRI; cdna made by oligo-dT priming.				
BASE COUNT	194 A 326 C 260 G 183 T				
ORIGIN					
Query Match	36.1%	Score 421.6:	DB 14:	Length 963:	
Best Local Similarity	78.5%:	Pred. NO. 1.2e+98:			
Matches 557;	Conservative 0:	Mismatches 144;	Indels 9;	Gaps 4:	
OY	193 AGTCGGCGCTCAACAAGAAGTGCTCTAAAGGCCGAAGCGCGCGCTCGCGCATATTGCG	252			
Dd	1 AGTTGGGCTCGCAAGAATGCGCACACTTAAGGCGCGA--AACGGGTCTGAAGACGATCCC	59			
OY	253 AGCCCATTTAGAGTTTCATCTCTGGSCAGAGCAGATGGAGGACAAAGCAGGTGGATGG	312			
Dd	60 AGCCCATTTAGAAATTCTATCCACGACCCTGACAGAGCAGGCGCAGCGAGGTGTGGACGG	119			
OY	313 GACACTGAGTGGCTGGGAAGACCAAAAATCAAAGCTTCACACCTCTGGCGTAGACCG	372			
Dd	120 GACACTGAGTGGCTGGGAAGACCAAGATAAAGCTTCACACCTCTGGCGTAGAACCG	179			
OY	373 CCAGATTGGGAATTTACAGTATCAGAGGCTGGCTTACTACTCTACTGTCAGGTGCA	432			
Dd	180 CCAGATTGGGAATTTATAGTACACCGGCGTGGCTTACTACTCTGTCAGGTGCA	239			
OY	433 CT TTGATAGAGGAAGAGCTGTCTACCTGAAGCTGTGACTTTGCTGGTGAACGAGTGTCTGGC	492			
Dd	240 CT TTGATAGAGGGAAGCGGTCTACTCTGAAGCTGACCTTGATGGATGAGTGTCTGGC	299			
OY	493 CCTGGCGTGGCTGGGAAGAAATTTCTAGAGCCACAGCAGCAAGCTTCCTGGGCCCCAGCTCG	552			
Dd	300 CCTGGCGTGGCTGGGAAGAAATTTCTAGAGCCACAGTGGGCCCAATTCCTTCGGGCCCCAGCTCG	359			
OY	553 TTTTGCCAAGTGTCTGGGCTTTTGCCGCTGGCGGACAGAGTTCCTTCCTTCGATCCGAC	612			
Dd	360 CCTTGCCAAGTGTCTGGGCTTTTGCCGCTGGCGGACAGAGTTCCTTCCTTCGATCCGAC	419			
OY	613 CCTCCCTGGGCTCACTTAAGAGTGCCTTCCTTAACCTACTTTGAGCTTTTCAAGT	672			

```

Db      420  COTCCCCCGGGCCCATCTCAAGGCTGGCCCCCTTCCTCAGCTACTTGGACTCTTCAGGT 479
QY      673  TCACGTAGAGGGCCCTTGGCTCTCCCGAGATCTCTTAACATTTCCCTGGCTCCAGAGACATCAC 732
Db      480  TCATCTGAAAGGGCCCTCGGTCTCCCAACAGTGTCTCCAGGCTGGCGGGCTCC-----CCTCGA 534
QY      733  CACACCTCCCTACGCCACACCCGACCTGCTCCAGCCGCCCTC-GCTGCTCTTGGTCCAGTCT 791
Db      535  CAGCTCTGTGGGACACCGCGTCCCTCTGCCCCACCTCAGCGCGCTCTTGTCTCAGAGCT 594
QY      792  GTCCTCTCC--TCAAAGGAGGCGACAGGTGTTCACATGTTTCATCTCCACAGAGCTATCC 849
Db      595  GCCCTCTCCCTTAGAGGGCTGCTGGGCGCTGTTCACGAGTGTTCATCTATCCCCATTAATACA 654
QY      850  TTGCTCTTCTTAACATCCGATCCACAGACAATATCCAGCTACACTAGCTC 899
Db      655  GTATTCACCACTCTATCTTACAACTCCGCCCAAGGCCAAGCTCTCAAGCTC 704

Search completed: March 31, 2003, 11:47:15
Job time : 2156.97 secs

```

Search completed: March 31, 2003, 11:47:15
Job time : 2156.97 secs

1	1168	100.0	1168	19	AAV1859
2	699.4	59.9	701	20	AAK23425
3	628.6	53.8	1353	21	AAA49717
4	628.6	53.8	1353	24	ABK40255
5	618.2	52.9	1421	20	AAK56000
6	618.2	52.9	1364	24	ABK34881
7	614.6	52.6	1373	19	AAV18600
8	597.8	51.2	1236	19	AAV47613
9	597.8	51.2	1236	22	AA004350

10	522.8	4.4	8	1030	20	AAK33424	Human TNFR1 DNA.
11	498.8	42.7	1093	22	AAK303964	Expression vector	
12	82.8	7.1	282	16	AAK722190	Human gene signal	
13	69.8	6.0	195	24	ABK929540	Colon adenocarcinoma	
14	65	5.6	65	24	ABN55975	Mouse spliced tran	
15	46.6	4.0	53522	24	AAAD30228	Human PKD1 gene.	
16	46.6	4.0	53526	19	AAAG94101	Human PKD1 gene.	
17	46.6	4.0	53577	17	AAAT18551	Human PKD1 gene.	
18	46.6	4.0	53577	14	AAAT94108	Human polytopic k	
19	46.6	3.9	105325	24	ABK94407	Human PKD1 locus b	
20	45	3.8	1337	20	AAAK217263	DNA encoding endob	
21	44.8	3.8	1000	21	AAAD20484	Human gene express	
22	43.6	3.7	16167	24	ABLT70254	Human colon cancer	
23	43.6	3.7	16167	24	ABLT33083	Chemically treated	
24	43.6	3.7	16167	24	ABLT4529	Human immune syste	
25	43.2	3.7	324	24	ABLT51826	Human metastasis a	
26	43.2	3.7	720	22	AAH05001	Human cDNA clone (
27	43.2	3.7	2260	22	AAH18456	Human DNA sequenc	
28	43.2	3.7	2272	22	AAH34689	Human DNA for a no	
29	43.2	3.7	2272	22	AAK34650	Human DNA for a no	
30	43.2	3.7	8604	24	ABOT1065	Listeria monocytog	
31	43	3.7	524	24	ABO34340	Oligonucleotide fo	
32	43	3.7	524	24	ABO34341	Oligonucleotide fo	
33	42.8	3.7	1166	20	AAZ19440	Oligonucleotide an	
34	42.8	3.7	1166	20	AAZ19228	M. tuberculosis an	
35	42.6	3.6	320	21	AAA38183	Primer used in the	
36	42.6	3.6	556	24	ABO36994	Oligonucleotide fo	
37	42.6	3.6	556	24	ABO36995	Oligonucleotide fo	
38	42.6	3.6	676	24	ABO36768	Oligonucleotide fo	
39	42.6	3.6	1286	24	ABO36769	Oligonucleotide fo	
40	42.4	3.6	1286	24	ABI99656	Mouse ischaemic co	
41	42.4	3.6	10732	21	AAAI0594	Gene encoding a su	
42	41.8	3.6	14006	24	ABL33959	Human immune syste	
43	41.6	3.6	1164	24	ABO68833	Listeria monocytog	
44	41.4	3.5	320	21	AAAS8186	Primer used in the	
45	41.4	3.5	1235	20	AAZ16147	Human gene express	

ALIGNMENTS

RESULT 1	
AAV18599	
ID	AAV18599 standard; cDNA; 1168 BP.

21-JUL-1998 (first entry)

DE Mus musculus tumour necrosis factor related ligand (TRELL) gene

KW TRELL; tumour necrosis factor related ligand; tnfr; treatment;
 KW cancer; autoimmune disease; immune system; stimulation; suppression
 KW graft rejection; ds.

OS Mus musculus

FH	Key	Location/Qualifiers
FT	CDS	2..679

/note= "tumour necrosis factor related ligand"

PN W09805783-A1

PD 12-FEB-1998

PF 07-AUG-1997; 97WO-US13945

PR 18-MAR-1997; 97US-0040820

PR	07-AUG-1990;	96US-0023341
PR	18-OCT-1996;	96US-0028515

XX
PA (BIOT) BIOGEN INC

PA

PA	(UUGE-)UNIV GENEVA FACULTY MEDICINE.
XX	
PI	Browning JL, Chicheportliche Y;
XX	
DR	WPI; 1998-145619/13.
XX	P-PSDB; AAM47524.
PT	Tumour necrosis factor related ligand - useful for, e.g. treating
PT	cancer, auto-immune disease and immune responses to tissue grafts
XX	
PS	Claim 2; Pages 45-46; 69pp; English.
XX	
CC	The sequence is that encoding mouse tumour necrosis factor related
CC	ligand (TRELL). TRELL or active fragments can be included with a
CC	carrier in pharmaceutical compositions to treat cancer, autoimmune
CC	diseases or immune responses to tissue grafts, or to stimulate or
CC	suppress the immune system. It is useful to screen for TRELL
CC	receptors, by labelling with a detectable label and screening
CC	compositions for binding. Agents interfering with TRELL-receptor
CC	binding can also be screened for, can then be administered,
CC	optionally with interferon-gamma, to induce cell death or
CC	treat, suppress or alter immune responses (especially involving human
CC	adenocarcinoma cells) involving a signal pathway between TRELL and its
CC	receptor. The DNA sequence can be used in gene therapy for
CC	TRELL-related disorders in mammals (especially humans), e.g. tumours,
CC	autoimmune and inflammatory diseases or inherited genetic disorders,
CC	by introducing into cells, and expressing, therapeutically effective
CC	amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
CC	It may also be of use in the preparation of prepare probes for
CC	screening natural/synthetic DNAs for TRELL-encoding sequences
CC	and for antisense therapy.
XX	
SQ	Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other:
	Query Match 100.0%; Score 1168; DB 19; Length 1168;
	Best Local Similarity 100.0%; Pred. No. 8 2e-307;
	Matches 1168; Conservative 0; Mismatches 0; Indels 0; Gaps
OY	1 GGTGCTGAGCCTGGGCGCTGGGCGTGCCTTGTCGCTGTCGTCAGCCT 60
Db	1 GGTGCTGAGCCGCGGCGCTGGGCGTGCCTTGTCGCTGTCGTCAGCCT 60
OY	61 GGGGAGCGTGGCAACGCTGTCGCCAGAGGCTTCTACGAGAGCTCACAGAGA 120
Db	61 GGGGAGCGTGGCAACGCTGTCGCCAGAGGCTTCTACGAGAGCTCACAGAGA 120
OY	121 CGCGCGGAGGCCCCCTGAATCATGCCAGACAGAGAAACCAGATGTGTACTTT 180
Db	121 CGCGCGGAGGCCCCCTGAATCATGCCAGACAGAGAAACCAGATGTGTACTTT 180
OY	181 CTGGACAACACTAGTCGGGCGCTGGAAGAAGTCTCTTAAGGCGGAAAGCGGCGCTG 240
Db	181 CTGGACAACACTAGTCGGGCGCTGGAAGAAGTCTCTTAAGGCGGAAAGCGGCGCTG 240
OY	241 CCGAECTATTGACGCCCATTTATGAGTTCATCTCGGCCAGACAKAGATGGACACAAC 300
Db	241 CCGAECTATTGACGCCCATTTATGAGTTCATCTCGGCCAGACAKAGATGGACACAAC 300
OY	301 AGGTGGATGGGACAGTAGAGTGGTGGGAAAGAACCAAATCAACAGCTCCAGGCTCT 360
Db	301 AGGTGGATGGGACAGTAGAGTGGTGGGAAAGAACCAAATCAACAGCTCCAGGCTCT 360
OY	361 GCAGTACAGACCGCCAGATTGGGGAATTTACATCAACAGAGCGCTGGCTTACTACTGTA 420
Db	361 GCAGTACAGACCGCCAGATTGGGGAATTTACATCAACAGAGCGCTGGCTTACTACTGTA 420
OY	421 CTGTACAGTGCATCTTGTATGAGGAAAAGCGTGTACCTGAAGCTGTGCTGGTAA 480
Db	421 CTGTACAGTGCATCTTGTATGAGGAAAAGCGTGTACCTGAAGCTGTGCTGGTAA 480
OY	481 CGGTGTGCTGGCGCTGGCGCTGCTGGGAAGATTTCTACAGCACAGACAAGCTTCTCG 540
Db	481 CGGTGTGCTGGCGCTGGCGCTGCTGGGAAGATTTCTACAGCACAGACAAGCTTCTCG 540

QY	541	GCCTCAGCTCCGTTTGTGGCCAGGAGTGTGGGGCTGTTGGCCGTGCGGACGAGGGCTCTTCCCT	600
Db	541	GCCTCAGCTCCGTTTGTGGCCAGGAGTGTGGGGCTGTTGGCCGTGCGGACGAGGGCTCTTCCCT	600
QY	601	TCGGATCCGCAACCTCCCTCGGGCTCATCTTAAAGCTGCGCCCTTTCCTTAACCTACTTTGG	660
Db	601	TCGGATCCGCAACCTCCCTCGGGCTCATCTTAAAGCTGCGCCCTTTCCTTAACCTACTTTGG	660
QY	661	ACTCTTTCAAGTTTACTGAGGGGCTTGGCTCTGCCAGATTCTTAACTTTTCCCTGGCTC	720
Db	661	ACTCTTTCAAGTTTACTGAGGGGCTTGGCTCTGCCAGATTCTTAACTTTTCCCTGGCTC	720
QY	721	CAGGAGATACACACACCTCCCTAACCCACACCCCTCCTCCACCCCTCGCTGCTCTT	780
Db	721	CAGGAGATACACACACCTCCCTAACCCACACCCCTCCTCCACCCCTCGCTGCTCTT	780
QY	781	GGTCCAGTCCCTGTCCTCTCTCAAAAGGACGACAGCTGTTGCATGTTCCATTCCACA	840
Db	781	GGTCCAGTCCCTGTCCTCTCTCAAAAGGACGACAGCTGTTGCATGTTCCATTCCACA	840
QY	841	GAGCTATCTTGGCTCTTCTTAAATCCATCCACACACAATATGCACTCACTAGCTCC	900
Db	841	GAGCTATCTTGGCTCTTCTTAAATCCATCCACACACAATATGCACTCACTAGCTCC	900
QY	901	CAAGAGCCCTACTATATCCCTGACTCCGCCACCCACTCACCCGACACAGTTTATTTGACT	960
Db	901	CAAGAGCCCTACTATATCCCTGACTCCGCCACCCACTCACCCGACACAGTTTATTTGACT	960
QY	961	TTGTGCACACAGGCACTGAGATGGGCTGGAGCTGGTGGCAGGAAGCCAGAACTGGGGAC	1020
Db	961	TTGTGCACACAGGCACTGAGATGGGCTGGAGCTGGTGGCAGGAAGCCAGAACTGGGGAC	1020
QY	1021	TAGGCCAAGATTTCCCAACTGTAGGGGGGAAGAGCTGGGGGACAAAGTCTCTCCCTGGATCC	1080
Db	1021	TAGGCCAAGATTTCCCAACTGTAGGGGGGAAGAGCTGGGGGACAAAGTCTCTCCCTGGATCC	1080
QY	1081	CTGTGATTTTGAAGAAAGATCTATTTTATTTATTTATTTGTAGCAAAATGTTAAATGGATAT	1140
Db	1081	CTGTGATTTTGAAGAAAGATCTATTTTATTTATTTATTTATTTGTAGCAAAATGTTAAATGGATAT	1140
QY	1141	TAAAGAGATTAATCATGATTTCTCTTC	1168
Db	1141	TAAAGAGATTAATCATGATTTCTCTTC	1168

PF 04-SEP-1998; 98MO-US18393.
 XX
 PR 05-SEP-1997; 97US-0924634.
 XX
 PA (UNIW) UNIV WASHINGTON.
 XX
 P1 Chaudhary PM;
 XX
 DR WP1: 1999-205191/17.
 DR P-PSDB; AAW93591.
 XX
 XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 XX
 PS Example VII, Fig 13B, 156pp; English.

CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: Ap04, Ap06, Ap08 and Ap09 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. Ap04 is useful for diagnosing prostate cancer
CC by determining levels of Ap04 in an individual. Prostate cancer can also
CC be treated using Ap04 selective binding agents linked to a therapeutic
CC moiety. Ap04 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed *in*
CC *in vivo*. Ap04 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the changer in Ap04
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using Ap04 polypeptides/active
CC fragments and Ap04 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of Ap04 and detecting a change in level of Ap04
CC activity. The method is performed *in vivo* or *in vitro*. Ap04 polypeptides
CC are all useful as immunogens for preparing antibodies. Ap04 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. Ap08 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.

XX Sequence 701BP; 139 A; 210 C; 203 G; 149 T; 0 other;

Query Match	59.9%	Score 699.4	DB 20	Length 701
Best Local Similarity	99.9%	Pred. No. 1.1e-179		
Matches 700	Conservative 0	Mismatches 1	Indels 0	Gaps 0
QY 44	CTGGTGTGTGTACGCTTGGGGAGCTGTGGGCAACGCTGTCTTGCCACAGAGCCTTCTCAGAG	103		
Db 1	CTGGTGTGTGTACGCTTGGGGAGCTGTGGGCAACGCTGTCTTGCCACAGAGCCTTCTCAGAG	60		
QY 104	GAGCTGCACGACAGAGAGGACCGCCGGGAGGCCCTGAACTGATGCCACAGAGGAAAGC	163		
Db 61	GAGCTGCACGACAGAGAGAGGACCGCCGGGAGGCCCTGAACTGAAATCCACAGAGGAAAGC	120		
QY 164	CAGATGTGTGTACCTTTCTTTGGCAACACTAGTCGGGCCCGGAAGAAAGTCTCTAAAGCC	223		
Db 121	CAGATGTGTGTACCTTTCTTTGGCAACACTAGTCGGGCCCTGGAAGAAAGTCTCTTAAAGCC	180		
QY 224	CGGAAGGCGCGGCTCTGCGCGAGCTATTGACGCCATTATGAGTTCATCTCTGGCCAGGA	283		
Db 181	CGGAAGGCGCGGCTCTGCGCGAGCTATTGACGCCATTATGAGTTCATCTCTGGCCAGGA	240		
QY 284	CAGATGTGAGCAACAAAGCAGTGTGGATGTGGATGGGACAGTGAATGGCTGGGAAAGACCAAAATC	343		
Db 241	CAGATGTGAGCAACAAAGCAGTGTGGATGGGACAGTGAATGGCTGGGAAAGACCAAAATC	300		
QY 344	AACAGCTTCACGCCCTCTCCGCTACACACCGCCAGATGTGGGAAATTTACATCATCAGGGCT	403		
Db 301	AACAGCTTCACGCCCTCTCCGCTACACACCGCCAGATGTGGGAAATTTACATCATCAGGGCT	360		
QY 404	GGGCTCTACTACTCTGTACTGTACAGTGCACCTTGATGAGGGAAGGCTGTCTACTGAAG	463		
Db 361	GGGCTCTACTACTCTGTACTGTACAGTGCACCTTGATGAGGGAAGGCTGTCTACTGAAG	420		

OY	464	CTGGACCTGCTGGGGAACGGGTGCTGGCCCTGGGCTGGAGAATTTCTCAGCCACA	523
OY	464	CTGGACCTGCTGGGGAACGGGTGCTGGCCCTGGGCTGGAGAATTTCTCAGCCACA	523
Db	421	CTGGACCTGCTGGGGAACGGGTGCTGGCCCTGGGCTGGAGAATTTCTCAGCCACA	480
OY	524	GCACGACGCTCTCCCTGGGGCCCAAGCTCCGTTTGTGCGACAGTGCTGGGCTGTGGCCGTG	583
Db	481	GCAGGACGCTCTCCCTGGGGCCCAAGCTCCGTTTGTGCGACAGTGCTGGGCTGTGGCCGTG	540
OY	584	CGGCCAGGCTCTCCCTGGGATTCGCGATCCGACCCCTCCCTGGGCTCATCTTAAAGCTGCCCC	643
Db	541	CGGCCAGGCTCTCCCTGGGATTCGCGATCCGACCCCTCCCTGGGCTCATCTTAAAGCTGCCCC	600
OY	644	TTCCGACCTCTTTGGAGCTCTTTCAGAGTTCACGGAAGGGGCTTGGCTCCGAGATTCCT	703
Db	601	TTCCGACCTCTTTGGAGCTCTTTCAGAGTTCACGGAAGGGGCTTGGCTCTCCGAGATTCCT	660
OY	704	TAAACTTTCCTGGCTCCAGAGCATCCACACACCTCCCTA	744
Db	661	TAAACTTTCCTGGCTCCAGAGCATCCACACACCTCCCTA	701
RESULT 3			
ID	AAAA9717		
XX	AAAA9717	standard; cDNA; 1353 BP.	
AC	AAA9717:		
XX			
DT	25-SEP-2000	(first entry)	
XX			
DE	Human PRO207	cdNA clone DNA30879-1152.	
XX			
KM	PRO207;	human; antitumour; tumour; therapy; cytostatic;	
KM	breast cancer;	ovarian cancer; renal cancer; colorectal cancer;	
KW	uterine cancer;	prostate cancer; lung cancer; bladder cancer;	
KW	central nervous system cancer;	melanoma; leukaemia; neoplasm; ss.	
XX			
OS	Homo sapiens.		
XX			
FH	Key	Location/Qualifiers	
FT	CDS	58..807	
FT		/*tag= a	
FT	sig_peptide	58..177	
FT		/*tag= b	
FT	mat_peptide	178..804	
FT		/*tag= c	
XX			
PN	MO200037638-A2.		
XX			
PD	29-JUN-2000.		
XX			
PE	02-DEC-1999;	99MO-US28565.	
XX			
PR	22-DEC-1998;	98US-0113296.	
PR	08-MAR-1999;	99MO-US05028.	
PR	21-APR-1999;	99US-0130232.	
PR	28-APR-1999;	99US-0131445.	
PR	14-MAY-1999;	99US-0134287.	
PR	20-JUL-1999;	99US-0144758.	
PR	26-JUL-1999;	99US-0145698.	
PR	15-SEP-1999;	99MO-US21090.	
PR	15-SEP-1999;	99MO-US21547.	
XX			
PA	(GETH)	GENENTECH INC.	
XX			
PI	Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA,		
PI	Napier MA, Pilti RM, Wood WI;		
XX			
DR	WPI: 2000-442668/38.		
DR	P-PSDB: AAY95338.		
XX			
PT	Novel composition to inhibit neoplastic cell growth or for treating		
PT	tumour in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO219,		
PT	PRO221, PRO224, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or		

PT PRO866 -
 XX Claim 20: Fig 3; 172pp; English.
 XX
 CC The present sequence is that of cDNA clone DNA30879-1152
 CC (ATCC 209338) encoding human PRO207 (see AAY95338), which shows
 CC homology to several members of the tumour necrosis factor family,
 CC especially human lymphotoxin (23.48). The cDNA was identified in a
 CC foetal kidney cDNA library following identification of an expressed
 CC sequence tag with homology to human Apo-2 ligand. A claimed method
 CC for inhibiting the growth of a tumour cell comprises exposing the
 CC tumour cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,
 CC PRO38, PRO301, PRO526, PRO362, PRO356, PRO509 or PRO866 (see
 CC AAY95337-49), their agonists or chimeric polypeptides incorporating
 CC them. The tumour is especially a cancer selected from breast,
 CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
 CC central nervous system cancer, melanoma and leukaemia. Nucleic
 CC acids encoding PRO179 etc. are used in the recombinant production
 CC of the antitumour polypeptides.
 XX
 SO Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other:
 Query Match 53.8%; Score 628.6; DB 21; Length 1353;
 Best Local Similarity 76.2%; Pred. No. 2,5e-160;
 Matches 933; Conservative 0; Mismatches 219; Indels 73; Gaps 10;

QY 722 AGGAGCATCACACACCTCCCTACCCACCACCTCTCCACACCCCTC-6GTGCTCTT 780
 DB 850 -----CTGCACAGCTCTCTGGGACACCGGCTCCCTCTGCCCCACCTCAGCGCTTT 904
 QY 781 GGTCCAGTCCCTGTCTCC-TCMAAGCAGCCAGAGCTTTGTACATGTTCCATTC- 837
 DB 905 GCTCCAGACCTGCCCTCCCTTAGAGGCTGCTGGGCTGTTCACGTGTTTCCATCCC 964
 QY 838 -----ACAGAGTATCTTGTCTTACATCTCCATCCACCACTATCCACCTC 891
 DB 965 ACATAAATACAGTATTCACACTCTTATCTTACAACTCCCAAGCCACTCTCACCTC 1024
 QY 892 ACTAGTCCCAAGCCCTCA-----TTATCCCTGACTCCCAACCCACT 936
 DB 1025 ACTAGTCCCAATTCCTGACCTTTGAGGCCCCCAAGTATCTGCACTCCCTGGCCA 1084
 QY 937 CACCCGACCACTGTTTATTTGACTTTGTGAC----- 968
 DB 1085 CAGACCCCGAGTATGTTTCACTGTACTCTGTGGCAGAGATGGTCCAGAGACCC 1144
 QY 969 -----CAGGACTGATAGGCTGACCTGTGGCAGGAAGCCAGAACCTGGGACTAG 1023
 DB 1145 CACTTACGACCTAAGAGGGGCTGACCTGGCGGAGGAAGCCAAAGACTGGGCTTAG 1204
 QY 1024 GCCAGAGTTCCCACTGTGTAGGGGAGAGCTGGGAGCAAGCTCCTCGTA-----TC 1079
 DB 1205 GCCAGAGTTCCCAATGTGTAGGGGAGAG-AACAAGACAGCTCCTCCTTGAAGATTG 1263
 QY 1080 CCTGTGATTTTGA--AGATCTATTTTATTTATTTGTACAAATGT---TAAAT 1134
 DB 1264 CCTGTGATTTTGAAGATATTTATTTATTTATTTGTACAAATGTGATTAAT 1323
 QY 1135 GATATTAAGATATAATCATGA 1159
 DB 1324 GATATTAATAGATAGTATTA 1348

RESULT 4
 ABK40255
 ID ABK40255 standard; cDNA: 1353 BP.
 XX
 AC ABK40255;
 XX
 DT 15-JUL-2002 (first entry)
 XX
 DE cDNA encoding human PRO207 polypeptide.
 XX
 KW Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
 KW Leukemia; neuronal disorder; stromal disorder; blastocellular disorder;
 KW Inflammatory disorder; immune disorder; angiolethic disorder;
 KW gene therapy; cytostatic; neuroprotective; gene; ss.
 OS Homo sapiens.
 XX
 PN WO200153486-A1.
 PD 26-JUL-2001.
 PF 11-FEB-2000; 2000MO-US03565.
 XX
 PE 08-MAR-1999; 99MO-US05028.
 PR 11-MAR-1999; 99US-123972P.
 PR 11-MAY-1999; 99US-133459P.
 PR 02-JUN-1999; 99MO-US12252.
 PR 22-JUN-1999; 99US-140653P.
 PR 22-JUN-1999; 99US-140653P.
 PR 20-JUL-1999; 99US-144758P.
 PR 26-JUL-1999; 99US-145698P.
 PR 28-JUL-1999; 99US-146222P.
 PR 31-AUG-1999; 99US-149395P.
 PR 31-AUG-1999; 99US-151689P.
 PR 01-SEP-1999; 99MO-US20111.

PN	US6207642-B1.
XX	
PD	27-MAR-2001.
XX	
PE	26-JUN-1998; 98US-0105343.
XX	
PR	12-FEB-1997; 97US-0798692.
PR	10-FEB-1998; 98US-0021706.
XX	
PA	(ABBO) ABBOTT LAB.
XX	
P1	Wiley SR;
XX	
DR	WPI: 2001-280760/29.
DR	P-Psdb: AA000891.
XX	
PT	Inducing angiogenesis in mammal at desired sites for promoting wound healing, by administering soluble fragment of extracellular domain of tumor necrosis factor related endothelium proliferative agent protein
PT	-
XX	
PS	Example 2; Column 73-74; 53pp; English.
XX	
CC	The present invention relates to extracellular signal molecules,
CC	particularly members of tumour necrosis factor (TNF) family molecules designated as TREPA (TNF related endothelium proliferative agent).
CC	Soluble biologically active TREPA are used to treat TREPA-associated diseases, tumours or metastases. TREPA is used for inducing angiogenesis in human for promoting wound healing and for vascularising grafted tissue for successful grafting and to promote tissue grafts.
CC	The present sequence is a cDNA clone ID #69050 encoding human TREPA.
XX	
SQ	Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;
Query Match	51.2%; Score 597.8; DB 22; Length 1236;
Best Local Similarity	75.2%; Pred No. 5.6e-152;
Matches	879; Conservative 0; Mismatches 222; Indels 66; Gaps
OY	2 GTGTGAGCGCTGGGCGCTGGGCGCTTGGCCTTCCTGCTGGTGCTACCGTG 61
DB	73 GCGCTGGGCGCTGGGCGCTGGGCGCTTCCTCGGCGCTCCTGCGCTGCATTTTG 132
OY	62 GGGAGCTGGGGCAACGCTGTCGCCAGAGACCCTTCACAGAGAGCTGACAGCAGAGAC 121
DB	133 GGGAGCCGGGCAATCGCTGTCCGCCAGAGACCTGCCACAGAGAGACTGTGGAGAGAG 192
OY	122 CGCGGGAGACCCCCTGAACCTGAATCCCCAGACAGAGAGAAAGCCAGAGTGTGTAACCTTC 181
DB	193 GACCGAGACCCGTCGGAACTGAATCCCCAGACAGAGAGAAAGCCAGAGTGTGCGCTTTC 252
OY	182 TTGGAAACAATACTAGTCCGGCTCGAAGAATGCTCCTTAAAGGCCGGAAGGCCGCGCTCGC 241
DB	253 CTGAACCCGACTGTGGTGGGCTCGCAAGAACTGCAACCTTAAAGGCCGGAAGAACACGCGCTTCA 312
OY	242 CGAGCATATTGCAAGCCATATGAGGTTCATCTCGGCCGACAGAGATGAGAGCAACAAGA 301
DB	313 AAGAGCATGCGACCCCATATTGAAGTCAATCCACAGACTCGACAGAGACGAGCGCAGGCA 372
OY	302 GGTTGTGATGGACACTGATGAGTGGCTGGGAAGAGACCAATAATCAACAGCTCCAGCCTCTG 361
DB	373 GGTGTGTGAGAGGGAAGTAGTGGCTGGGAGGAAGCCAGATCAACAGCTCCAGCCTCTG 432
OY	362 GCCTAGAGACCGGCAAGTTGGGGAAATTACAGTCATCAGAGGCGGGCTACTACTCCGTGAC 421
DB	433 CCCTACAAACCGCAAGTAGTCGGGGAGTTTAATGTCACCCGGGCTGAGCTTCACTACGTGAC 492
OY	422 TGTCAGGTGCTACTTTGATGATGAGGAAAAGCGCTGTTACTACTGAGCTGACTTGTGCTGAG 481
DB	493 TGTCAGGTGCTACTTTGATGATGAGGGAAGGCGTGTACTGTAGCTGTGACCTTGCGTGGAGAT 552
OY	482 GGTGTGTCTGGCCCTGCGTGCCTGGAGAAGATTCTCACGCCACAGCAAGCTTCTCTGGG 541
DB	553 GGTGTGTCTGGCCCTGCGTGCCTGGAGAAGATTCTCACGCCACAGCAAGCTTCTCTGGG 612

QY	542	CCGAGCTCCGTTTgTgCCAGAGGTGTCGGGCGTTCGCGCCAGAGGCTTCCTT	601
QY	542	CCGAGCTCCGTTTgTgCCAGAGGTGTCGGGCGTTCGCGCCAGAGGCTTCCTT	601
Db	613	CCCCAGCTCCGGCTCTgCCAGGtGTCTTGGGCTGTGTGGCTTCGGGCTCGGCTC	672
QY	602	CGAGTTCGCGACCCCTCCCTGGGCTCATTTAAAGGCTGCCCTTCCTTAACCTATTGGA	661
Db	673	CGAGTTCGCGACCCCTCCCTGGGCTCATTTAAAGGCTGCCCTTCCTTAACCTATTGGA	732
QY	662	CTCTTTCAAGTTCAGTGAAGGGGCTTGTCTCCAGATTCTTTAAACTTTTCCCTGGCTC	721
Db	733	CTCTTTCAAGTTCAGTGAAGGGGCTTGTCTCCAGATTCTTTAAACTTTTCCCTGGCTC	792
QY	722	AGGAGCATTCACACACCTCCCTACCCACCCGACCTCTGCACCCGCTC-GCTGCTCTT	780
Db	793	-----CCTCGACAGCTCTTCTGGGACCCGGTCCCTCTCTGCCACGCTCAGCGGCTCTTT	847
QY	781	GGTCAGTTCCTGTCTCTCC--TCAAAGGCGAGCCAGAGCTTGTTCACATGTTTCATTCC- 837	
Db	848	GCTCCAGACCGGCCCTCCCTCTTGAAGGGCTGCTGGGCTGTTCACAGTGTTTTCCATCCC 907	
QY	838	-----ACAGACGATCTCTTGTCTTCTTTAACATCCCATCCACACAGCTATCCACTC 891	
Db	908	ACATAAATACAGTATCTCCACTCTTATCTTACAAACAACCCAGCGCCACTCTCCACCTC 967	
QY	892	ACTGAGCTCCCAAGCCCTCTAC-----TTATTCCTGACTCCCCACCCACTC 936	
Db	968	ACTGAGCTCCCAATCTCTGACCCCTTTGAGGCCCCACAGTATCTCGACTCTCCCTCGGCCA 1027	
QY	937	CACCCGACCAAGTGTATTGACTTTGTGTGCAC----- 968	
Db	1028	CAGACCCCGAGGGGACTGTGTTCACGTACTGTGTGGCAAGATGGGTCAGAAAGCCC 1087	
QY	969	-----CAGGCACGTGAATGGGCTGTGACCTGTGTGGCAGAGCAAGCTGGGACTAG 1023	
Db	1088	CACCTTCAGGCACTAAGAGGGGGCTGTGACCTGTGGCAGAGCAAGCAAGACTGSGCTTAG 1147	
QY	1024	GCCAGAGTTCCTCCAACTGTAAGGGGGAGAGGCTGGGACCAAGCTCCTCCCTGGA----TC 1079	
Db	1148	GCCAGAGTTCCTCCAACTGTAAGGGGGAGAG--AACAGACAAAGCTCCTCCCTTGAGAAATTC 1206	
QY	1080	CCTGTGATTTTGAAGAAAGTACTATTTTT 1108	
Db	1207	CCTGTGATTTTAAACAGATATATTTTT 1235	
RESULT 10			
AAAX23424			
XX	AAAX23424	standard; DNA: 1030 BP.	
XX	AAAX23424:		
XX	18-JUN-1999	(first entry)	
XX	Human TNRL3 DNA.		
XX	Tumour necrosis factor receptor; signal transducer molecule; TNF; AP04;		
KW	developmental abnormality; gestational abnormality; prostate cancer;		
KW	AP06; AP08; AP09; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;		
KW	cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;		
KW	apoptosis; human; ss.		
OS	Homo sapiens.		
XX	Key	Location/Qualifiers	
PH	CDS	1..627	
FT		/*tag= a	
FT		/product= "TNRL3"	
XX	MO9911791-A2.		
XX	11-MAR-1999.		
XX			

PF 04-SEP-1998; 98BW-US18393.
 XX
 PR 05-SEP-1997; 97US-0924634.
 XX
 PA (UNIV) UNIV WASHINGTON.
 XX
 PI Chaudhary PM;
 XX
 DR WPI; 1999-205191/17.
 DR P-PSDB; AAW93590.
 XX
 XX New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 XX
 PS Example VII; Fig 13A; 156pp; English.

This invention describes isolated Tumor Necrosis Factor (TNF) family receptor polypeptides: AP04, AP06, AP08 and AP09 or their active fragments, and isolated TNF related ligands 1 and 3 (TNFR1 and TNFR3) or their active fragments. AP04 is useful for diagnosing prostate cancer, by determining levels of AP04 in an individual. Prostate cancer can also be treated using AP04 selective binding agents linked to a therapeutic moiety. AP04 polypeptides are also useful for identifying selective binding agents, useful in diagnosis/treatment of disease by binding of agents to the polypeptide/active fragment which is extracellular, or expressed on the cell surface. The binding is preferably performed in vivo. AP04 polypeptides/active fragments are also useful for screening for agonists and antagonists by binding and observing the changer in AP04 activity. Effective pharmacological agents useful in diagnosis or treatment of disease are also identified using AP04 polypeptides/active fragments and AP04 signal transducer molecules that specifically interact with a cytoplasmic domain of AP04 and detecting a change in level of AP04 activity. The method is performed in vivo or in vitro. AP0 polypeptides are all useful as immunogens for preparing antibodies. AP0 polypeptides are useful for diagnosis/treatment of developmental or gestational abnormalities. AP08 was transfected to human breast carcinoma cell line MCF-7, and induced apoptosis.

SQ Sequence 1030 BP; 223 A; 317 C; 279 G; 211 T; 0 other;

Query Match	44.88	Score 522.8	DB 20	Length 1030
Best Local Similarity	81.68	Pred. No. 1.2e-131		
Matches 643; Conservative	0	Mismatches 137	Indels 8	Gaps 3

QY	53	GTAGCCTGGGGAGACTGGGGACAGCCTGTCTGCCAGAGAGCCTTCTCAGAGAGAGCTGACA	112
Db	1	GTCAGTTTTGGGAGCCGGGAGCTGCTCTCGGCCAGAGAGCCTGCCAGAGAGAGCTGGTG	60
QY	113	GCAAGAGACCGCCGGAGAGCCCTCGAAGCTGAATCCCGACAGAGAGAAACCAAGATGTG	172
Db	61	GCAAGAGAGAGACAGAGACCCGTGGAACTGAATCCCGACAGAGAGAAACCCAGATCTT	120
QY	173	GTACCTTTCTTGGAACTAGTCCGGCCTCGAAGAGAGCTCCTTAAGGCCGGAGGCG	232
Db	121	GGCGCTTTCGAAACCCAGTATGTTGGGCTCTCGAGAGTGCACCTAAAGCCGGAAAAACA	180
QY	233	CGGCTCGCCGAGCTATTTGCAAGCCCATTTATAGAGTTATCTGGGCCAGAGACAGATGA	292
Db	181	CGGGCTCGAAGAGGAGATCGCAGCCCATTAATGAATTCATCCAGACCTGGACAGGAGGA	240
QY	293	GCACAAACAGAGTGTGATGTGGAGACAGTGAAGGCTGGGAGAGACCAAAATCAACAGCTCC	352
Db	241	GGCAGGAGAGTGTGTGACAGGGACAGTGAATGAGTGTGGGAGAGACCAAAATCAACAGCTCC	300
QY	353	AGCCCTGTGGCTACGACACCGCCAGATTGGGGAATTTACAGTCAATCAGAGGCTGGGCTTAC	412
Db	301	AGCCCTGTGGCTACCGCCAGATGGGAGATTATAGACACCGCGGCTGGGCTTAC	360
QY	413	TACCTGTACTGTCAAGGTGCACTTTGATGAGAGGAAAGCCTGTACTCTGAAGCTGGACTTG	472
Db	361	TACCTGTACTGTCAAGGTGCACTTTGATGAGAGGAAAGCCTGTACTCTGAAGCTGGACTTG	420

QY	473	CTGGGAACGGGTGTGCTGGGCCCTGCGGCTGTGGAAATTTCTCAGCCACAGCAGCAAGC	5322
Db	421	CTGGGAGTGGTGTGTGCTGGGCCCTCGCTGCGCTGTGGAGAAATTTCTAGCCACAGCGGGCAGT	4800
QY	533	TCTCTCTGGGGCCCAAGCTCCCTTTTGTGCAGAGTGTCTGGGGCTGTTTGGCCGTGGCGGCCAGGG	5922
Db	481	TCCCTCGGGGGCCCCAGCTCCCGCTCTGCCAGGTGTCTGGGGCTGTTGGGCCCTTGGCGGCCAGGG	5400
QY	593	TCTTCCCTTCGGATCCGACCCCTCCCTGGGGCTCATCTTAAGGCTGCCCCCTTCTTAAC	6522
Db	541	TCTTCCCTTCGGATCCGACCCCTCCCTGGGGCCATCATCAAGGCTGCCCCCTTCTTAAC	6000
QY	653	TACTTTGGACACTTTTCAAGTTCACTAGAGGGGCTTGTCTTCCCAAGATTTCTTTAACTTTC	7122
Db	601	TACTTTGGACACTTTTCAAGTTCACTAGAGGGGCTTGTCTTCCCAAGATTTCTTTAACTTTC	6600
QY	713	CTGTGCTCAGAGAGATCACACACACTTCCCTTAACCCCAACCCCACTCTCCACACCCCTC-G	7722
Db	661	GCGGGCTCC-----CTTGCACAGCTCTCTTGTGGGACACCGGGTCCCTTGTGCCCAACCTTCAG	7150
QY	772	CTGTGCTCTGTGCTCAGTCTGTCTCTCC--TCAAAAGCAGGACAGAGCTTGTTCATCATGTT	8250
Db	716	CCGGCTTGTGCTCAGACACTGCCCTCCCTCTAGAGGGTGGCTGGGCGCTGTTCACAGTGT	7750
QY	830	TCCATTCG 837	
Db	776	TTCCATTCG 783	

RESULT 11
AAS03964

ID AAS03964 standard; DNA; 898 BP.

AC AAS03964 ;

DT 26-SEP-2001 (first entry)

DE Expression vector pDC409-LZ-TWEAK fusion protein-encoding DNA

KM TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;
KM ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
KM retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
KM ruberosis; uveitis; macular degeneration; arthritis; rheumatism; ds;
KM corneal graft neovascularisation; psoriasis; metastatic condition;
KM malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
KM preneoplastic condition; myocardial angiogenesis; wound granulation;
KM scleroderma; vascular adhesion; telangiectasis; ischaemia; human;
KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;
KM peripheral atherosclerosis; pdc409-L2-TWEAK; TWEAK receptor; TWEAKR;
KM fusion protein.

05 'Homo sapiens.
05 Synthetic.

FH	Key	Location/Qualifiers
FT	CDS	52..873

FT	/product=	"Fusion protein comprising a growth hormone
FT		leader, a leucine zipper multimerisation
FT		domain, and human TWEAK extracellular
FT		domain"

PN W0200145730-A2

PD 28-JUN-2001

PF 19-DEC-2000; 2000WO-US34755.

PR 20-DEC-1999; 99US-0172878.

PR 10-MAY-2000; 2000US-0203347.

PA (IMMV) IMMUNEX CORP.

XX

PI Wiley SR:
XX
DR WPI: 2001-417975/44.
DR P-PSDB: AAU03499.
XX
PT Modulating angiogenesis in a mammal for treating diseases mediated by
PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
PT peripheral tissue, by administering antagonist or agonist of TWEAK
PT receptor
XX
PS Example 1; Page 39-40; 46pp; English.
XX
CC The sequence represents a DNA from the expression vector
CC pPC409-12-TWEAK, which encodes a fusion protein comprising a growth
CC hormone leader, a leucine zipper multimerisation domain, and the
CC extracellular domain of human TWEAK. The fusion protein was used in
CC the isolation of human TWEAK receptor (TWEAKR)-expressing clones
CC from a COS cell human cDNA library. The TWEAK protein is a
CC member of the tumour necrosis factor (TNF) family and induces
CC angiogenesis. TWEAKR may therefore be used to screen for and develop
CC TWEAKR agonists and antagonists for the modulation of angiogenesis, to be
CC used in the treatment and diagnosis of human disease. The disorders
CC mediated by angiogenesis include ocular disorders characterised by ocular
CC neovascularisation such as diabetic retinopathy, neovascular glaucoma,
CC retinoblastoma, retinopathy of prematurity, retrolental fibroplasia,
CC rubeosis, uveitis, macular degeneration and corneal graft
CC neovascularisation, and inflammatory diseases such as arthritis,
CC rheumatism and psoriasis. Other treatable diseases include malignant and
CC metastatic conditions such as sarcomas and carcinomas, benign tumours and
CC preneoplastic conditions, myocardial angiogenesis, haemophilic joints,
CC scleroderma, vascular adhesions, atherosclerotic plaque
CC neovascularisation, telangiectasia, wound granulation, coronary
CC atherosclerosis, peripheral atherosclerosis and ischaemia.
XX
SQ Sequence 898 BP; 187 A; 266 C; 267 G; 178 T; 0 other:

Query Match 42.7%; Score 498.8; DB 22; Length 898;
Best Local Similarity 87.0%; Pred. No. 3.6e-125;
Matches 548; Conservative 0; Mismatches 82; Indels 0; Gaps 0;

QY 56 AGCTGGGAGAGCTGGGAGAGCGTGTGCGGAGAGCGCTTTCAGAGAGCTGACAGCA 115
DB 250 AGTTGGGAGAGCGGCGATCGCTCTCGCCAGAGAGCTGCCAGAGAGAGCTGTGCA 309
QY 116 GAGGAGCCGCGGAGAGCGGAGAGTGAATCCAGAGAGAGAGAGAGAGAGAGTGTGTA 175
DB 310 GAGGAG 369
QY 176 CCTTCTTGGAG 235
DB 370 CCTTCTTGGAG 429
QY 236 CCTGCGGAG 295
DB 430 GCTGAG 489
QY 296 CAAGCAGAGTGTGAG 355
DB 490 CAGCAGAGTGTGAG 549
QY 356 CCTTGGCGTGTGAG 415
DB 550 CCTTGGCGTGTGAG 609
QY 416 CTGTACTGTGAG 475
DB 610 CTGTACTGTGAG 669
QY 476 GTGAAGAGTGTGAG 535
DB 670 GTGAGTGTGAG 729
QY 536 CTGTGGGCGGAG 595

DB 730 CTGGGCGGAG 789
QY 596 TCCCTTGGAGATCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 655
DB 790 TCCCTTGGAGATCCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 849
QY 656 TTTGAGCTCTTCAAGTTCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 685
DB 850 TTGGAGCTCTTCAAGTTCAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 879

RESULT 12
AAT22190
ID AAT22190 standard; cDNA to mRNA; 282 BP.
XX
AC AAT22190;
XX
DT 27-AUG-1996 (first entry)
XX
DE Human gene signature HUMG503761.
XX
KW Gene signature; messenger RNA; mRNA; relative abundance; frequency;
KW human; cloning; mapping; non-biased library; diagnosis; detection;
KW cell typing; abnormal cell function; ss.
XX
OS Homo sapiens.
XX
PN W09514772-A1.
XX
PD 01-JUN-1995.
XX
PF 11-NOV-1994; 94WO-JP01916.
XX
PR 12-NOV-1993; 93JP-0355504.
XX
PA (MATS/) MATSUBARA K.
PA (OKUB/) OKUBO K.
PI Matsubara K, Okubo K;
XX
DR WPI: 1995-206931/27.
XX
PT Identifying gene signatures in 3'-directed human cDNA library - e.g.
PT for diagnosis of abnormal cell function, by preparing cDNA that
PT reflects relative abundance of corresp. mRNA in specific human
PT tissues
XX
PS Claim 1; Page 1067; 2245pp; Japanese.
XX
CC A single-stranded DNA (or its complementary strand or the corresp.
CC double-stranded DNA) which comprises one of the 7837 "GS" sequences
CC given in AAT19001-726837 and which is able to hybridise to one of
CC human genomic DNA, cDNA or mRNA is claimed. The GS (Gene signature)
CC sequences were obtained from 3'-directed cDNA libraries prepared
CC from various human tissues; synthesis of cDNA was initiated from the
CC 3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-
CC untranslated sequence is unique to a particular mRNA species, almost
CC all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library
CC is constructed so as to reflect accurately the relative abundance of
CC different mRNAs in the particular tissue from which it was derived.
CC The appearance frequency of a given GS in a cDNA library can be
CC determined (esp. using primers and probes derived from the GS
CC sequences) as a means of diagnosing abnormal cell function or for
CC recognising different cell types.
XX
SQ Sequence 282 BP; 80 A; 62 C; 69 G; 66 T; 5 other:

Query Match 7.1%; Score 82.8; DB 16; Length 282;
Best Local Similarity 80.6%; Pred. No. 1.9e-12;
Matches 158; Conservative 0; Mismatches 26; Indels 12; Gaps 5;

QY 969 CAGGCACTGAGATGGGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1027

||||| ||| ||||||||| | ||||||||| ||| ||||| |||||||
Db 88 CAGGCACTAAGAGGGGCTGACCTNTGCGGAGGAAGCAAGACTGGGCTAGGCCA 147
Oy 1028 GAAGTTCCTCAACTGTGAGGGGGAAGAGCTGGGACAGCTCTCCCTGGA----TCCCTG 1083
Db 148 GGAGTTCCTCAACTGTGAGGGGGAAGAGCTGGGACAGCTCTCCCTGGA----TCCCTG 206
Oy 1084 TGGATTTCGAAA--AGATACATATTTATTTATTTATTTGACAAAATGTT----AAATGGA 1137
Db 207 TGGATTTTAAAACAGATATTTATTTTNTATTTATTTGACAAAATGTTGNTAAATGGGA 266
Oy 1138 TATTAAGAGAAATAA 1153
Db 267 TATTAATAGAAATAA 282
RESULT 13
ABK29540
ID ABK29540 standard; cDNA: 195 bp.
AC ABK29540;
XX 23-APR-2002 (first entry)
Df
XX Colon adenocarcinoma-specific cDNA #66.
DE Human: colon adenocarcinoma; colon cancer; tumour; gene; ss.
XX Homo sapiens.
OS WO200196389-A2.
XX 20-DEC-2001.
PD 07-JUN-2001; 2001WO-US18574.
PF 09-JUN-2000; 2000US-210667P.
XX 22-NOV-2000; 2000US-252614P.
PR (CORI-) CORIXA CORP.
XX (CORI-) CORIXA CORP.
PA Meagher MJ, King GE, Xu J, Secrist H;
PI WPI: 2002-098052/13.
DR
XX New isolated polynucleotide encoding a polypeptide comprising a portion
PT of colon tumour protein, for detection, diagnosis and therapy of human
PT colon cancer
PS Claim 1: Page 133; 211pp; English.
XX The invention relates to an isolated polynucleotide (I) encoding a
CC polypeptide (II) comprising a portion of a colon tumour protein. A new
CC oligonucleotide (III) that hybridises to (I) is useful for
CC determining the presence of a cancer in a patient. (II) or antigen
CC presenting cells expressing (I) are useful for stimulating and/or
CC expanding T cells specific for a tumour protein, by contacting T cells
CC with (I), (II) or antigen-presenting cells that express (I). (I), (II),
CC or antigen presenting cells that express (II) are useful for treating
CC colon cancer in a patient by incubating CD4+ and/or CD8+ T cells isolated
CC from a patient with (I), (II), or antigen presenting cells that express
CC (II), so that T cells proliferate, and administering to the patient an
CC effective amount of the proliferated T cells, thus inhibiting the
CC development of a cancer in the patient. A new composition is useful for
CC stimulating an immune response in a patient. (I) or (II) is useful in
CC vaccines and pharmaceutical compositions for prevention and treatment of
CC colon cancer and for the diagnosis and monitoring of the cancers. (I),
CC (II) or an antibody against (II) is useful for detection, diagnosis and/
CC or therapy of human colon cancer. (I) is useful as a probe or primer for
CC nucleic acid hybridisation, and in the design and preparation of ribozyme
CC molecules for inhibiting expression of (II) in tumour cells. ABK29475-
CC ABK29851 represent human colon adenocarcinoma-specific cDNA sequences of
CC the invention.

XX
SQ Sequence 195 bp; 49 A; 51 C; 58 G; 37 T; 0 other;
Query Match 6.0%; Score 69.8; DB 24; Length 195;
Best Local Similarity 83.5%; Pred. No. 5,4e-09;
Matches 91; Conservative 0; Mismatches 17; Indels 1; Gaps 1;
Oy 969 CAGGCACTAGATGGGCTGACCTGATGTCAGAGAGCCAGAGAACTGGAGCTAGGCCAG 1028
Db 69 CAGGCACTAAGAGGGGCTGACCTGGCCGACAGAGCCAAAGAGACTGGGCTAGGCCAG 128
Oy 1029 AAGTTCCTCAACTGTGAGGGGGAAGAGCTGGGACAGCTCTCCCTGGA 1077
Db 129 GAGTTCCTCAACTGTGAGGGGGAAGAGCTGGGACAGCTCTCCCTGGA 176
RESULT 14
ABN55975
ID ABN55975 standard; DNA: 65 bp.
AC ABN55975;
XX 15-JUL-2002 (first entry)
Df
XX Mouse spliced transcript detection oligonucleotide SEQ ID NO:28723.
DE Human: mouse; rat; splice transcript; detection; RNA transcript;
XX splice variant; transcriptome; oligonucleotide library; ss.
KM Mus musculus.
OS WO200210449-A2.
XX 07-FEB-2002.
PD 20-JUL-2001; 2001WO-IB01903.
PF 28-JUL-2000; 2000US-221607P.
XX 02-MAY-2001; 2001US-287724P.
PR (COMP-) COMPUGEN INC.
XX Shoshan A, Wasserman A, Mintz E, Mintz L, Faigler S;
PI WPI: 2002-257383/30.
DR
XX New oligonucleotide libraries comprising oligonucleotides which
PT selectively hybridize to mRNAs transcribed from a transcription unit of
PT a genome, useful for detecting tissue-, pathology-, and
PT developmental-specific genes
PS Example 1: SEQ ID 28723; 47pp; English.
XX The present invention describes oligonucleotide libraries for detecting
CC messenger RNAs that populate a (sub-)transcriptome, where the
CC (sub-)transcriptome comprises messenger RNAs transcribed from multiple
CC transcription units that populate a genome. The library comprises
CC several oligonucleotides, each capable of hybridising selectively to a
CC set of messenger RNAs transcribed from a given transcription unit of
CC the genome, which encodes one or more messenger RNA splice variants.
CC The oligonucleotide libraries are useful for detecting mRNAs from a
CC biological sample, in expression profiling studies, in qualitatively or
CC quantitatively characterising the corresponding transcriptome, and in
CC detecting RNA transcripts and splice variants of human or animal
CC transcriptomes. The libraries may also be used as specialised mini
CC libraries to detect transcripts of a sub-transcriptome under a
CC particular biological or pathological state, and so allowing the
CC detection of tissue- and pathology-specific genes such as those genes
CC only expressed in specific tissue under a specific pathological
CC condition; to detect developmental specific genes; and to detect RNA
CC transcripts and splice variants of a transcriptome of a patient suffering
CC from a particular disorder. ABN27253 to ABN59589 represent
CC oligonucleotide sequences from rats, humans and mice, which are used in

CC the exemplification of the present invention.
CC N.B. The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 65 BP; 14 A; 15 C; 21 G; 15 T; 0 other;

Query Match 5.6%; Score 65; DB 24; Length 65;
Best Local Similarity 100.0%; Pred. No. 6.5e-08;
Matches 65; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1030 AGTTCCTCACTGTGAGGGGAGAGCTGGGACAACTCCTCCCTGGATTCCTGTGATT 1089

DB 1 AGTTCCTCACTGTGAGGGGAGAGCTGGGACAACTCCTCCCTGGATTCCTGTGATT 60

QY 1090 TTGAA 1094

DB 61 TTGAA 65

RESULT 15

AAD30228
ID AAD30228 standard; DNA: 53522 BP.

XX AAD30228;

DT 17-MAY-2002 (first entry)

XX Human PKD1 gene.

XX Human: PKD1 gene; autosomal dominant polycystic kidney disease; ADPKD;
KW acquired cystic disease; transgenic animal; Chromosome 16; ds.

XX Homo sapiens.

OS W0200206529-A2.

PN 24-JAN-2002.

PD 13-JUL-2001; 2001WO-US22035.

PF 13-JUL-2000; 2000US-218261P.

PR 13-APR-2001; 2001US-283691P.

XX (UYJO) UNIV JOHNS HOPKINS SCHOOL MEDICINE.

PA Germino GG, Matnick TJ, Phakdeekitcharoen B;

PI WPI; 2002-179805/23.

DR Novel primer for diagnosing polycystic kidney disease-associated
XX disorder, comprises regions having sequence that selectively hybridizes
PT to polycystic kidney disease gene sequence -

XX Claim 20; Page 127-156; 192pp; English.

PS The present invention relates to compositions and methods useful for the

CC identification and detection of polycystic kidney disease (PKD1) gene
CC mutations. The invention also relates to primers comprising a 5' region

CC having a sequence that selectively hybridizes to a PKD1 gene sequence
CC and optionally, to a PKD1 homologue sequence and an adjacent 3' region

CC having a sequence that selectively hybridizes to a PKD1 gene sequence
CC and not to a PKD1 homologue sequence. Primer pairs of the invention are

CC useful for detecting the presence or absence of a mutation in a PKD1
CC polynucleotide in a sample, for identifying a subject at risk for a

CC PKD1-associated disorder such as autosomal dominant polycystic kidney
CC disease (ADPKD) or acquired cystic disease and for diagnosing a PKD1-

CC associated disorder in a subject. They are useful for selectively
CC amplifying a region of a PKD1 gene. PKD1 DNA fragments are useful

CC detecting the presence of a mutant PKD1 polynucleotide in a sample,
CC as a probe for an amplification reaction, in hybridisation or

CC amplification assays of biological samples to detect abnormalities
CC of PKD1 expression and for engineering transgenic animals. The present

CC sequence is human PKD1 gene located on chromosome 16.

XX Sequence 53522 BP; 8486 A; 17664 C; 15766 G; 11606 T; 0 other;

QY Query Match 4.0%; Score 46.6; DB 24; Length 53522;
Best Local Similarity 48.7%; Pred. No. 0.15;
Matches 127; Conservative 0; Mismatches 134; Indels 0; Gaps 0;

QY 685 CTTCCTCTCCAGATTCTTAACCTTTCCTGGCTCCAGAGCATACACACCTCCCTA 744

DB 34365 CTTCCCTCTCCCTACCCCTTCCCTCTCCCTCCCTAGACCTTCCCTACCTCTCC 34424

QY 745 CCCCACCCCACTCTCCACCCCTCTGCTCTCTTGTCAAGTCTGTCTCTCAAA 804

DB 34425 CGCTGAGACCCCTCCACTGCTGCTCCCAAGCCCTCCCTAGACCCCTCCCTCCCTT 34484

QY 805 GGCAGCCAGAGCTTGTTCACATGTTTCATTTCCACAGACGATGCTCTTCTTACA 864

DB 34485 CTTCCCTCTCTCCCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 34544

QY 865 TCCCATCCACACACATATCAGCTTCCCAAGCCCTTACTTATCCCTGACT 924

DB 34545 TTCTCTCCCT 34604

QY 925 CCCCACCCCACTCTACCCGACC 945

DB 34605 TCTCTCCCT 34625

Search completed: March 31, 2003, 08:16:43
Job time : 431.94 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: March 31, 2003, 07:58:15 ; Search time 3683.27 Seconds

(without alignments)
9228.782 Million cell updates/sec

Title: US-09-245-198a-1

Perfect score: 1168

Sequence: 1 ggtgcgagagctcgctcg.....ataatcatgattctcttc 1168

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_vt:*

31: em_htg_hum:*

32: em_htg_inv:*

33: em_htg_other:*

34: em_htg_mus:*

35: em_htg_pln:*

36: em_htg_rtd:*

37: em_htg_mam:*

38: em_htg_vrt:*

39: em_sy:*

40: em_htgo_hum:*

41: em_htgo_mus:*

42: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1168	100.0	1168	10	AF030100 Mus muscu
2	711	60.9	203083	2	AC069459 Mus muscu
3	711	60.9	234182	10	AL603707 Mouse DNA
4	628.6	53.8	1353	6	AX201324 Sequence
5	628.6	53.8	1368	6	AF055872 Homo sapi
6	624	53.4	1306	9	AF030099 Homo sapi
7	597.8	51.2	1236	6	AR140407 Sequence
8	501	42.9	138792	2	AC119115 Rattus no
9	498.8	42.7	1651	9	AX180714 Sequence
10	409.4	35.1	1651	9	BC019047 Homo sapi
11	304	26.0	215795	2	AC127470 Pan trogl
12	303.8	26.0	177703	2	AC016876 Homo sapi
13	278.2	23.8	153553	2	AC126921 Bos tauru
14	237.4	20.3	177555	2	AC130192 Sus scrof
15	212	18.2	161428	2	AC126925 Canis fam
16	156.8	13.4	184026	2	AC098923 Rattus no
17	111.6	9.6	203281	2	AC126237 Canis fam
18	88.4	7.6	7218	6	166494 Sequence 14
19	86.4	7.4	139405	2	AC126239 Felis cat
20	69.8	6.0	1195	6	AX379024 Sequence
21	61	5.2	190376	2	AC123372 Rattus no
22	58.4	5.0	64424	2	AC124103 Mus muscu
23	56.2	4.8	114260	2	AC098956 Rattus no
24	56.2	4.8	125020	9	AF429315 Homo sapi
25	55.8	4.8	80112	2	AC128800 Rattus no
26	55.4	4.7	69119	2	AC128983 Rattus no
27	55.4	4.7	241370	2	AC094065 Rattus no
28	55	4.7	185822	2	AC073554 Homo sapi
29	54.8	4.7	100511	2	AC010774 Homo sapi
30	54.6	4.7	221758	2	AC068947 Mus muscu
31	54.4	4.7	87120	2	AC012225 Homo sapi
32	54.2	4.6	298166	2	AC087563 Homo sapi
33	54	4.6	111627	2	AC108626 Rattus no
34	54	4.6	184036	2	AC126109 Rattus no
35	54	4.6	310371	2	AC096296 Rattus no
36	53.8	4.6	99517	2	AC131402 Rattus no
37	53.6	4.6	205350	2	AC078946 Mus muscu
38	53.4	4.6	146986	2	AC095995 Rattus no
39	53.4	4.6	203281	2	AC126237 Canis fam
40	53.2	4.6	49430	2	AC100434 Mus muscu
41	53.2	4.6	182269	2	AC107416 Rattus no
42	53	4.5	55061	2	AC091597 Mus muscu
43	53	4.5	176645	2	AC114115 Rattus no
44	53	4.5	187252	10	AL607109 Mouse DNA
45	52.8	4.5	84514	10	AL627264 Mouse DNA

ALIGNMENTS

RESULT 1

AF030100

LOCUS AF030100 1168 bp mRNA 1linear ROD 20-DEC-1997

DEFINITION Mus musculus TWEAK mRNA, partial cds.

ACCESSION AF030100

VERSION AF030100.1 GI:2707220

KEYWORDS

SOURCE

ORGANISM Mus musculus.

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclerognathhi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 1168)

Chicheportliche,Y., Bourdon,P.R., Xu,H., Hsu,Y.M., Scott,H., Hession,C., Garcia,I. and Browning,J.L.

TWEAK, a new secreted ligand in the tumor necrosis factor family

JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL
that weakly induces apoptosis	J Biol Chem. 272 (51), 32401-32410 (1997)	98070415	2 (bases 1 to 1168)	Chicheportiche,Y., Blixier,S., Tizard,R. and Browning,J.	Direct Submission	Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA
FEATURES	source	1..1168	/organism="Mus musculus"	/db.xref="taxon:10090"	/cell.type="peritoneal macrophages"	<1..679 /note="Ligand in the TNF family; secreted protein" /codon_start=2 /product="TWEAK" /protein_id="AAC53517.1" /db_xref="GI:2707221" /translation="VLSIGTLALAGLLLVYSISGNMTSLASQPSDELTAEDRREF PELNPTEESODVPFLDLQLAPRRRSAPKGRAPRIIAHYEHPPGDDGAOGAVY DGVTSWGEETKINSSPLRYDRDIGFTVIYAGLYUYCOVHPDEGRAVKYLKLIVN GYALRLGLEERSATNAASSPQQLRLQVSGILPIRPSSLRIRLPMAHLKAAPFLTYVH" FGLEOVH"
CDS						
BASE COUNT	242 a	360 c	298 g	268 t		
ORIGIN						
Query Match	Best Local Similarity	100.0%; Score 1168;	DB 10;	Length 1168;		
Matches 1168;	Conservative	0;	Mismatches	0;	Indels	0;
Gy	1	GGTGCTAGACCTCGGGCCCTGGCGCTGCGCTGCCTTCCTGCTGCTGTGTACGCT	60			
Dd	1	GGTGCTAGACCTCGGGCCCTGGCGCTGCGCTGCCTTCCTGCTGCTGTGTACGCT	60			
Gy	61	GGGAGACTGGGCAACGCTGTCTGCCAGAGAGCCTCTCAGAGAGAGCTACAGAGAGA	120			
Dd	61	GGGAGACTGGGCAACGCTGTCTGCCAGAGAGCCTCTCAGAGAGAGCTACAGAGAGA	120			
Gy	121	CGCGCGGAGCCCCCTGAACTGTAATCCCAGACAGAGAAAAGCCAGATGTGTACTTT	180			
Dd	121	CGCGCGGAGCCCCCTGAACTGTAATCCCAGACAGAGAAAAGCCAGATGTGTACTTT	180			
Gy	181	CTTGAACAACACTAGTCCGGCCCTCGAAGAACTGCTCTTAAAGCCGGAAGCCGGCTCG	240			
Dd	181	CTTGAACAACACTAGTCCGGCCCTCGAAGAACTGCTCTTAAAGCCGGAAGCCGGCTCG	240			
Gy	241	CGGACTATTGTCAGCGCATTTATGAGTTTCATCTTCGGGCGACAGACAGATGGACCAAGC	300			
Dd	241	CGGACTATTGTCAGCGCATTTATGAGTTTCATCTTCGGGCGACAGACAGATGGACCAAGC	300			
Gy	301	AGGTGTGATGGGAACAGTAGAGTGGTGGGAAGAGACAAAAATCAACAGCTCCAGCCCTCT	360			
Dd	301	AGGTGTGATGGGAACAGTAGAGTGGTGGGAAGAGACAAAAATCAACAGCTCCAGCCCTCT	360			
Gy	361	GGCGTACGACCGCCAGATTGGGGAAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTA	420			
Dd	361	GGCGTACGACCGCCAGATTGGGGAAATTTACAGTCATCAGGGCTGGGCTCTACTACCTGTA	420			
Gy	421	CTGTAGAGTGCACCTTTGATGAGGAGAAAGGCTGTCTACTGTAAGCTGAGCTTCTGTGTAA	480			
Dd	421	CTGTAGAGTGCACCTTTGATGAGGAGAAAGGCTGTCTACTGTAAGCTGAGCTTCTGTGTAA	480			
Gy	481	CGGTGTGTCGGGCCCTGGCGTGCCTGGGAAGAAATTCACGACACAGAGCAAGCTCTCTCGG	540			
Dd	481	CGGTGTGTCGGGCCCTGGCGTGCCTGGGAAGAAATTCACGACACAGAGCAAGCTCTCTCGG	540			
Gy	541	GCCCCAGCTCGCTTTGTGTGCAAGTGTCTGTGGCGCTGTGGCCGACAGGGTCTTTCCT	600			
Dd	541	GCCCCAGCTCGCTTTGTGTGCAAGTGTCTGTGGCGCTGTGGCCGACAGGGTCTTTCCT	600			
Gy	601	TGCGATCCGACACCCTCCCTGGGCTATCTTAAGGCTGCCCCCTTCTAACCTTACTTTGG	660			

D6	601	TGCGATCCGGCACACCTCCCTGGGCTCATTCTTAAGGCTGGCCCCCTTCCTAACTACTGTTGG	660
OY	661	ACTCTTTCAGATTCACTAGAGGGGCTTGCTGCCAGATTCTCTTAACATTTTCCGTGGCTC	720
D6	661	ACTCTTTCAGATTCACTAGAGGGGCTTGCTGCCAGATTCTCTTAACATTTTCCGTGGCTC	720
OY	721	CAGGAGCATMCCACACACCTCCCTACCACCCCACTCCCTCCACCCCGCTGGTCCCTT	780
D6	721	CAGGAGCATMCCACACACCTCCCTACCACCCCACTCCCTCCACCCCGCTGGTCCCTT	780
OY	781	GGTCCAGTCTGTCTCTCTCTCAAAGGACGACGACCTGTGCATATGTTTCCATTCCACA	840
D6	781	GGTCCAGTCTGTCTCTCTCTCAAAGGACGACGACCTGTGCATATGTTTCCATTCCACA	840
OY	841	GACGTATCTCTGTCTCTCTCTTAACATCCATCCACCAACAATATCCACTCACTAGCTCC	900
D6	841	GACGTATCTCTGTCTCTCTCTTAACATCCATCCACCAACAATATCCACTCACTAGCTCC	900
OY	901	CAAAACCCCCTACTTATATCCCTGACTCCCCACCCACTCACCGACACGATTCACCTCAGCTGC	960
D6	901	CAAAACCCCCTACTTATATCCCTGACTCCCCACCCACTCACCGACACGATTCACCTCAGCTGC	960
OY	961	TGTGTCCACGACGACATGAGATGGCGCTGACCTGTGTGACGAGAAGCCAGAGAACCTGGGAC	1020
D6	961	TGTGTCCACGACGACATGAGATGGCGCTGACCTGTGTGACGAGAAGCCAGAGAACCTGGGAC	1020
OY	1021	TAGGCCAGAACTTCCCAACTGTGAGGGGAGAGAGCTGGGGACAACTCTCCTCTGATCC	1080
D6	1021	TAGGCCAGAACTTCCCAACTGTGAGGGGAGAGAGCTGGGGACAACTCTCCTCTGATCC	1080
OY	1081	CTGTGGATTTTGAAGAAGTACTATTTTTATTTATTTATGTGACAAAATGTTAATAGATAT	1140
D6	1081	CTGTGGATTTTGAAGAAGTACTATTTTTATTTATTTATGTGACAAAATGTTAATAGATAT	1140
OY	1141	TAAAGAGATAAATCATGATTTCTCTTTC	1168
D6	1141	TAAAGAGATAAATCATGATTTCTCTTTC	1168
RESULT 2			
LOCUS	AC069459/C		
DEFINITION	Mus musculus chromosome 11 clone RP23-168P5, WORKING DRAFT		
ACCESSION	AC069459		
VERSION	AC069459.23 GI:14547768		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE	Mus musculus.		
ORGANISM	Eutheria; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
REFERENCE	1 (bases 1 to 203083)		
AUTHORS	Weizker,M.L., Lewis,L.R., Hume,J., Edwards,C., Harris,C., Dedering,D., Thomas,S., Okunomu,G., Carlock,C., Garner,T., Addison,S., Pace,A., Williams,G., Bonnin,D., Brooks,A., Brown,J., Bunay,C., Bunnac,C., Burkett,C., Chacko,J., Chen,G., Chen,Z., Cox,C., Davis,C., Delgado,O., Ding,Y., Dugan-Rocha,S., Fernandez,C., Ferraguto,D., Forcum-Tansey,J., Gill,R., Correll,J.H., Gunaratne,P., Haller,G., Hernandez,J., Hogues,M., Hosak,H., Hou,X., Huber,J., Jackson,L., Jia,Y., Kelly,J., Kelly,S., Kovar,C., Liu,J., Liu,W., Louissege,H., Lozado,R.J., Martin,R., Massey,E., McLeod,M.P., Mel,G., Moore,S., Morgan,M., Morris,S., Neal,D., Nelson,A., Nguyen,R., Ogih,N., Oguh,M., Parish,B., Perez,L., Relter,D., Say,J., Shen,H., Vasquez,L., Watlington,S., Williamson,A., Wrensford,G., Zhou,X., Bouck,J., Hodgson,A., Muzny,D.M., Rives,M., Scherer,S., Sodergren,E., Wetnstock,G., Worley,K.K. and Gibbs,R.		
JOURNAL	Direct Submission		
REFERENCE	2 (bases 1 to 203083)		
AUTHORS	Worley,K.C.		
TITLE	Direct Submission		

In the feature table with their source databases: Em., EMBL, Sw., SWISSPROT, Tr., TREMBL, Wp., WORMPEP, information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-422L16 is from the RPI-23 Mouse PAC Library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm> VECTOR: pBAC3.6

This sequence is the entire insert of clone RP23-422L16.

FEATURES

source
1..234182
Location/Qualifiers
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone.lib="RPI-23"
complement(84050..84131)
/note="Sequence from uni-directional primer reads and dGTP
big dye terminator reads only."
misc_feature
59310 a 56824 c 57519 g 60529 t
BASE COUNT
ORIGIN

Query Match 60.9%; Score 711; DB 10; Length 234182;
Best Local Similarity 99.6%; Pred. No. 2.6e-176;
Matches 744; Conservative 0; Mismatches 0; Indels 3; Gaps 3;

OY 425 CAGGTGACCTTTGATGAGGAAGCTGTCTACCTGAGCTGACCTTGGTGAACGGT 484
|||||
Db 6623 CAGGTGACCTTTGATGAGGAAGCTGTCTACCTGAGCTGACCTTGGTGAACGGT 66264
OY 485 GTGCTGGCCCTGGCTGCTGGAAGAAATCTCAGCCACAGACAGCAAGCTCTCTGGGCC 544
|||||
Db 66263 GTGCTGGCCCTGGCTGCTGGAAGAAATCTCAGCCACAGACAGCAAGCTCTCTGGGCC 66204
OY 545 CAGCTCGTTTGGCCAGAGTGTCTGCTGCTTCCCGCTGGGCCAGAGGCTTCCCTTGG 604
|||||
Db 66203 CAGCTCGTTTGGCCAGAGTGTCTGCTGCTTCCCGCTGGGCCAGAGGCTTCCCTTGG 66144
OY 605 ATCCGACACCTCCCTGGGCTCATCTTAAAGCTGCCCCCTTCAACCTACTTGGACCT 664
|||||
Db 66143 ATCCGACACCTCCCTGGGCTCATCTTAAAGCTGCCCCCTTCAACCTACTTGGACCT 66084
OY 665 TTTCAAGTTCTAGTGAAGGGGCTTGTCTCCAGATTCTTAACTTTCCTGCTCCAG 724
|||||
Db 66083 TTTCAAGTTCTAGTGAAGGGGCTTGTCTCCAGATTCTTAACTTTCCTGCTCCAG 66024
OY 725 AGATACACACACCTCCCTGACCCACCCACCTCCACCTCCCTGCTGCTGCTGCTG 784
|||||
Db 66023 AGATACACACACCTCCCTGACCCACCCACCTCCACCTCCCTGCTGCTGCTGCTG 65964
OY 785 CAGTCCCTGT-CTCTCTCAAGAGCAGCAGAGCTTGTTCAGATG-TTTCATTCCACAGA 842
65963 CAGTCCCTGTCTCTCTCAAGAGCAGCAGAGCTTGTTCAGATGTTTTCATTCCACAGA 65904
OY 843 CGTATCTTGTCTCTTC-TTAACTCCATCCACACCAACTATCCACTAGCTGCC 901
|||||
Db 65903 CGTATCTTGTCTCTTCCTTAACTCCATCCACACCAACTATCCACTAGCTGCC 65844
OY 902 AAGCCCTTATCTATCCCTGACCTCCCAACCCACTACCCGACGACGCTGTTTATGACT 961
65843 AAGCCCTTATCTATCCCTGACCTCCCAACCCACTACCCGACGACGCTGTTTATGACT 65784
OY 962 TGTGACACAGGACAGTGAAGCTGAGCTGAGCTGAGCAGAGCAAGCAAGCTGGAGCT 1021
65783 TGTGACACAGGACAGTGAAGCTGAGCTGAGCTGAGCAGAGCAAGCAAGCTGGAGCT 65724
OY 1022 AGGCAAGAACTTCCCACTGTGAAGGGGAGAGCTGGGGAACAAGCTCTCCGATGCC 1081
65723 AGGCAAGAACTTCCCACTGTGAAGGGGAGAGCTGGGGAACAAGCTCTCCGATGCC 65664
OY 1082 TGTGAGTTTGAAGAACTATATTTTATATATATGTGCAAAATGTTAAATGATAT 1141
65663 TGTGAGTTTGAAGAACTATATTTTATATATATGTGCAAAATGTTAAATGATAT 65604

OY 1142 AAGAGATAATCATGATTTCTCTTC 1168
Db 65603 AAGAGATAATCATGATTTCTCTTC 65577

RESULT 4
LOCUS AX201324 1353 bp DNA linear PAT 30-AUG-2001
DEFINITION Sequence 3 from Patent WO0153486.
ACCESSION AX201324
VERSION AX201324.1 GI:15391154
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE
1 (bases 1 to 1353)
Ashkenazi, A.J., Goddard, A., Godowski, P.J., Gurney, A.L.,
Hillan, K.J., Marsters, S.A., Pan, J., Pitt, R.M., Roy, M.A., Smith, V.,
Stone, D.M., Watanabe, C.K. and Wood, W.I.
Compositions and methods for the treatment of tumour
Patent: WO 0153486-A 3 26-JUL-2001;

TITLE
JOURNAL
Genentech, Inc. (US)

FEATURES

source
1..1353
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
BASE COUNT 257 a 443 c 389 g 264 t
ORIGIN

Query Match 53.8%; Score 628.6; DB 6; Length 1353;
Best Local Similarity 76.2%; Pred. No. 7.5e-155;
Matches 933; Conservative 0; Mismatches 219; Indels 73; Gaps 10;

OY 2 GTGCTAGCCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 61
Db 130 GCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 189
OY 62 GGGAGTGGGCAAGGCTGTCTGCGCAGAGGCTTCTCAGAGAGCTGACAGAGAGAG 121
Db 190 GGGAGTGGGCAAGGCTGTCTGCGCAGAGGCTTCTCAGAGAGCTGACAGAGAGAG 249
OY 122 CGCGGAGAGCCCTGACCTGATGATCCACAGACAGAGAGAGAGAGAGAGAGAGAG 181
Db 250 GACGAGAGCCCTGACCTGATGATCCACAGAGAGAGAGAGAGAGAGAGAGAGAGAG 309
OY 182 TTGGAACAACTAGTCCGGCTGAGAGAGTCTCTTAAAGGCGGAGAGGCGGCTGCC 241
Db 310 CTGAAACCACTAGTCTGGGCTGACAGAGTGCACCTTAAAGGCGGAGAGAGAGAGAG 369
OY 242 CGAGCTATTGACGCCATTTATGAGTTTCTGCTGCGCAGAGAGAGAGAGAGAGAG 301
Db 370 AGAGCAATGACAGCCATTTATGAGTTTCTGCTGCGCAGAGAGAGAGAGAGAGAGAG 429
OY 302 GGTGTGATGGGACAGTGAAGTGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 361
Db 430 GGTGTGATGGGACAGTGAAGTGGCTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 489
OY 362 CGCTACGAGCCGCAAGATTGGGGAATTACAGTATATAGGGCTGGGCTGACTACCTGAC 421
Db 490 CGCTACGAGCCGCAAGATTGGGGAATTATAGTACCCGGGCTGGGCTGACTACCTGAC 549
OY 422 TGTGACGACCTTTGATGAGGAGAGAGGCTGTACTGAGCTGAGCTGAGCTGAGAC 481
Db 550 TGTGACGACCTTTGATGAGGAGAGAGGCTGTACTGAGCTGAGCTGAGCTGAGAC 609
OY 482 GGTGTGCTGGCCCTGGCTGCTGGAAGAAATCTCAGCCACAGACAGCAAGCTCTCTGG 541
Db 610 GGTGTGCTGGCCCTGGCTGCTGGAAGAAATCTCAGCCACAGCTGAGGAGGCTCTGG 669
OY 542 CCCGACCTCGTTTGTGCAAGGTGTGTGGGCTGTCGCGCTGGGCGGACAGGCTTCTCC 601
|||||

Db	670	CCCGAGCTCCGCTCTGGCAGGTCGTGGGGCTGTGGCCCTGGGGCCAGAGGCTCCCTCG	729
Oy	602	CGGATCCGACCCCTCCCTGGGGCTCATCTTAAAGCTGCCCCCTTCTTAACCTACTTTGGA	661
Db	730	CGGATCCGACCCCTCCCTGGGGCCCATCTCAAGGCTGGCCCCCTTCTCACTACTTCGGA	789
Oy	662	CTCTTCAAGTTTCACTGAGGGGCTTGGCTCCGCCAGATTCTTAACATTTCCTGGGGCTC	721
Db	790	CTCTTCCAGGTTCACTGAGGGGCTCTGGCTTCGCCGCAAGTCGTCCAGGCTGCGGGCTCC	849
Oy	722	AGGAGCATCACACACACCTCCATCCCAACCCCACTCCCTCCACCCCTTC - GCTGCTCCTT	780
Db	850	-----CGTCGACAGCTCTCTGGGACCGGGGTCCCTCTGCCCAACCTCAGCGGCTCTTT	904
Oy	781	GGTCACAGTCTGCTCTCTCC - TCAAGGACGACGAGCTTGTTCATGTTTCCATTC -	837
Db	905	GCTCCAGACCTGCGCCCTCTCTAAGAGCTGCTGGGCTGTTCACGTTTTTCCATCCC	964
Oy	838	-----ACAGACGTAATCCCTTGTCTTCTTAACATCCCAATCCCAACAACTATCCACCTC	891
Db	965	ACATATAATACGATATTCCCACTTATCTTACACTCCCCCAACGCGCCACTCTCCACTC	1024
Oy	892	ACTAGCTCCCAAGGCCCTTAC - -----TTATCCCTGATCTCCCAACCCACT	936
Db	1025	ACTACTCTCCCAATCCCGACCCCTTTGAGGGCCCCAGATCTGCATCTCCCCCTGGGCA	1084
Oy	937	CACCCGACCACTGTATTATGTGACTTTGTGCAC - -----	968
Db	1085	CAGACCCCCAGCTGTTGTTCACCTGTACCTCTGTGGCAAGATGGGTCCAGAGACCC	1144
Oy	969	-----CAGGACCTGATAGGCTGTGACCTGTGTGGAGAGAACCCAGAACTGGGACTAG	1023
Db	1145	CACCTCAGGCACTAAGAGGGGCTGGACCTGGGGGAGAGAACCAAGAGACTGGGCTTAG	1204
Oy	1024	GCCAGAACTTCCCACTGTGAGGGGAGAGAGCTGGGAGCAAGCTCTCCCTGGA - --TC	1079
Db	1205	GCGAGGATTTCCCAATGTGAGGGGCGGAGA - AACAAGCAAAAGCTCTCTTGAGAAATTC	1263
Oy	1080	CTGTGTGATTTTGA - -AGATACTATTTTATTTATTTATTTGTGACAAATGT - --TAAAT	1134
Db	1264	CTGTGTGATTTTGA - -AGATACTATTTTATTTATTTATTTATTTGTGACAAATGT	1323
Oy	1135	GGATTTTAAAGAGATAATCATGA	1159
Db	1324	GGATTTTAAAGATAATCATGA	1348
RESULT 5			
AF05872			
LOCUS	AF05872	1368 bp	mRNA
DEFINITION	Homo sapiens Ap03/DR3 ligand (Ap03L) mRNA, complete cds.		linear PRI 04-MAY-1998
ACCESSION	AF05872		
VERSION	AF05872.1	GI:3108230	
KEYWORDS			
SOURCE			
ORGANISM			
	Homo sapiens.		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
REFERENCE			
AUTHORS	Maesters,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and Ashkenazi,A.		
TITLE	Identification of a ligand for the death-domain-containing receptor Ap03		
JOURNAL	Curr. Biol.	8 (9), 525-528	(1998)
MEDLINE	982828355		
PUBMED	9560343		
REFERENCE	2 (bases 1 to 1368)		
AUTHORS	Maesters,S.A., Sheridan,J.P., Pitti,R.M., Brush,J., Goddard,A. and Ashkenazi,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-MAR-1998)		
FEATURES	South San Francisco, CA 94080, USA		
	Location/Qualifiers		

[illegible]

OY	838	----	-ACAGAGAGTATTCCTGCTCTCTTAACATCCATCCACCAAGAAATATCCACCTC	891
Db	925	ACATAAATACAGATATATCCACTCTTATCTTACAACTCCCCACGCCCACTCTCAACCTC	984	
OY	892	ACTAGCTCCCAAGCCCCCTAC-----TTATCCCTACTCCCCACCACT	936	
Db	985	ACTAGCTCCCAATCCCTGACCCCTTTGAGGGCCCCAGATCTGCACTCCCCCTGGGCA	1044	
OY	937	CACCCGACCACTGTTTATTGACTTTGTCAC-----	968	
Db	1045	CAGACCCCCAGGGCACTTGTGTTCACTGTACTGTGGGCAAGATGGGTCCAGAAAGACC	1104	
OY	969	-----CAGGCACTGAGATGGGCTGGACCTGGTGGGAGAAAGCAACCTGGGACTAG	1023	
Db	1105	CACTTCCAGGCACTAAGAGAGGGCTGGACCTGGGGGCGAGAAAGCAAGACTGGGCCCTAG	1164	
OY	1024	GCCAGAAATTTCCCAACTGTGAGGGGGAAGAGACTGGGGGCAAGCTCCTCCCTGGA-----TC	1079	
Db	1165	GCCAGGATTTCCCAATGTGAGGGGCGAGA-AACAAGACAAAGCTCTCTCCCTTGGAAATTC	1223	
OY	1080	CCTGTGATTTTGA--AGATACTATTTTATTTATTTATTTGTGACAAAATGT--TAAT	1134	
Db	1224	CCTGTGATTTTAAAAACAGATATATTTATTTATTTATTTGTGACAAAATGTTGATAAT	1283	
OY	1135	GGATTTAAAGAGATAATCA	1156	
Db	1284	GGATTTAAATAGATAATGCA	1305	

RESULT 7			
ARI40407			
LOCUS	1336 bp	DNA	
DEFINITION	Sequence 1 from patent US 6207642.	linear	PAT 16-JUN-2001
DESCRIPTION			

SOURCE	ORGANISM
Unknown.	Unknown.
Unknown.	Unclassified.
REFERENCE	1 (bases 1 to 1236)

BASE COUNT	ORIGIN	FEATURES	JOURNAL	TITLE
225	a	416 c	358 g	237 t
				1. .1236
				Location/Qualifiers
				Patent: US 6207642-A 1 27-MAR-2001;
				Member of the TNF family useful for treatment and diagnosis of disease

Query Match	51.2%:	Score	597.8:	DB	6:	Length	1236:
Best Local Similarity	75.2%:	Pred. No.	1e-146:				
Matches	879;	Conservative	0;	Mismatches	222:	Indels	68;
						Gaps	8;

67 GAGCAGACCGGACCTGGCCATGACCCTGAGCCCTGAGTGTGGTCAGGCCG 61

73 GGCGTGGGCCTGGGCCCTGGCGCTGCCCTGCTCTGCTGGCCGTGGCAATTG 132

Db 133 GGGAGCCCCGGCTCGCTGTCCCGCCAGAGAGCTGCCCCAGGAGAGCTGTGTGGCAGAGAG 192

182 TTGGAACTAACTGTCTCCGCTTGAAGAGTGTCTCTTAAAGGCTGGAAAGCGTGGGCTTCC 241
193 GACACAGACCCCTCGGAACATGATCCCGACAGAAAGCAAGCCAGATCCTGTGGCCCTTTC 252
Db

253	CTGAAACCGACTAGTTCTGGCCCTCGAAGAAGTGCACCTTAAAGCCGGGAAACACGCGGCTCGA	312
Db		
242	CGAGCTATTGCAGCCCATTTATGAGTTCATCCCTCGGCCAGGACAGGATGAGGCACAAGCA	301
Oy		

Db 313 AGAGCGATCGCAGCCCCATTATGTAAGTTCATCCACGACCTGGACAGGACGGAGCGCAGAGCA 372

Oy	302	GGTGTGATGGGACAGTAGTGGCTGGGGAAGAGACCAAAATTCACAGCTCCAGCCCTGTG	361
Db	373	GGTGTGGACGGGACAGTAGTGGCTGGGGAAGGACCAAAATTCACAGCTCCAGCCCTTG	432
Oy	362	CGCTACGACCCGAGATTTGGGGAATTTACAGTCACAGGCGTGGGCTCTACTACTGTAC	421
Db	433	CGCTACACCCGCAAGATGGGGAGTTTATAGTCACCCGGGCTGTACTACTGTAC	492
Oy	422	TGTGAGTGACATTGGATGAGGGAAAGCGCTGTACTGTGAAGCTGGACTGTGGTGTAC	481
Db	493	TGTGAGTGACATTGGATGAGGGAAAGCGCTGTACTGTGAAGCTGGACTGTGGTGTAT	552
Oy	482	GATGAGTGAGCCCTGCGCTGTGCTGGAAATTTCTAGCAGCAGCAGCAAGCTCTCTGGG	541
Db	553	GATGAGTGAGCCCTGCGCTGTGCTGGAAATTTCTAGCAGCAGCAGCAAGTTCCCTCGGG	612
Oy	542	CCCGAGCTCCGTTGTGTGCGCAGGTGTCTGGGCTGTGTCCGCTGCGGCGCAGGGCTTTCCCTT	601
Db	613	CCCGAGCTCCGCTGTGCGCAGGTGTCTGGGCTGTGTCCGCTGCGGCGCAGGGCTTCCCTG	672
Oy	602	CGGATCGCAGCCCTCCCTGGGAGCTCATCTTAAGGCTGCGCCCTTCTCAACTACTTTGGA	661
Db	673	CGGATCGCAGCCCTCCCTGGGAGCTCATCTTAAGGCTGCGCCCTTCTCAACTACTTTGGA	732
Oy	662	CTCTTTCAAGTTCACTGAGGGGCGTTGCTCTGCCAGATTCCTTAAACTTTCCCTGCGTCC	721

Dy 722 AGGAGCATCACCACCACTCCCTTACCCCACCCCCCAATCCTCCACGCCCGTC - GCTGTCTTT 780
 | || ||| | ||| ||| ||| |||
Db 793 -----CTTCGACAGCTCTCTGGGCACCCGGTCCCTTGCCCGCACCCCTCAGCGGCTTTT 847

Oy 838 -----ACAGACGATCCCTTGCTGCTTTCTTAACATCCCATCCACCACCAACTATCCACCTC 891
 | | | | | | | | | |
Dd 908 ACATAAATACGATATTCCCATCTTATCTTTACAACACACCCACCGCCGCATCTTCACACTC 967

QY	852	ACTAGCTCCCAAGGCCCTAC-----TTATCCCTGACTCCCCACACCCT	936
Db	968	ACTAGCTCCCAATCCCTGACCCCTTGGGCCCCCGATGATCTGACTCCCCCTGGCCA	1027
QY	937	CACCCGACACAGTGTATTATGACTTTGGAC-----	968
Db	1028	CAGACCCCCAGGGCATTTGTTCACCTGTGACCTGTGGCGAAGATGGGTCCAGAAGACC	1087

[illegible]

Db	1148	GGCAGGAGTTC	CCAAAAAGTGAAGGGCGAGA	-AACACGACAACCTCCCTCCGAGAAATTC	1206
QY	1080	CCGTGGATTTT	GAAAGAGATACATATTTT	1108	

Db 1207 CCTGTGCATTTTAAACAGATATTATT 1235

AC119115			
AC119115	138792 bp	DNA	linear
LOCUS			HTG 18-JUL-2002
Rattus norvegicus	clone CH230-320N23, ***	SEQUENCING IN PROGRESS	
DEFINITION	*** 32 unordered pieces		

ACCESSION
VERSION
KEYWORDS
SOURCE
AC119115.2
GI:21746718
HTG; HTGS_PHASE1.
Norway rat.

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae:
Rattus.
1 (bases 1 to 138792)

Muzny,D.M., Adams,C., Adio-oduola,B., Ali-osman,F.R., Allen,C.,
Alstbrooks,S.L., Amaralunge,H.C., Ayele,M., Ayele,M., Banks,T.,
Barbieri,J., Benton,J., Blincke,K., Blankenburg,K., Bonnin,D.,
Bouch,J., Bowles,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
Doutwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Correll,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hayes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B.,
Homs,F., Howard,S., Huber,J., Huiyk,S., Hume,J., Jackson,L.E.,
Jacobsen,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsso,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
Kralovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtenage,O., Lileu,C., Liu,J., Liu,W., Louisedge,H.,
Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Meador,M., Mei,G., Metzger,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokenkwo,S., Ogulu,M., Okwuonu,G.,
Oraguonye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
Rieves,M., Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Severy,G.,
Scherer,S., Scott,G., Shen,Y., Shooshari,N., Sisson,I.,
Sodergren,E., Sonake,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansley,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S.,
Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G., and Gibbs,R.

TITLE
JOURNAL
REFERENCE

Unpublished
2 (bases 1 to 138792)

Worley,K.C.

TITLE
JOURNAL
REFERENCE
Submitted (25-APR-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA

3 (bases 1 to 138792)

Worley,K.C.

REFERENCE
AUTHORS
JOURNAL
Submitted (18-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 14, 2002 this sequence version replaced gi:20303440.

COMMENT
Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

Project Information

Center project name: GING

Center clone name: CH230-320N23

Summary Statistics

Sequencing vector: Plasmid

Chemistry: Dye-terminator Big Dye: 100% of reads

Assembly program: Phrap: version 0.990329

Consensus quality: 112720 bases at least Q40

Consensus quality: 11666 bases at least Q30

Consensus quality: 119165 bases at least Q20

NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank/draft_data.html).
NOTE: This is a 'working draft' sequence. It currently
consists of 32 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.

1 1022: contig of 1022 bp in length
1023 1122: gap of unknown length
1123 2394: contig of 1272 bp in length
2395 2494: gap of unknown length
2495 3663: contig of 1169 bp in length
3664 3763: gap of unknown length
3764 4879: contig of 1116 bp in length
4880 4979: gap of unknown length
4980 6787: contig of 1808 bp in length
6788 6888: gap of unknown length
6888 8192: contig of 1305 bp in length
8193 8292: gap of unknown length
8293 9461: contig of 1169 bp in length
9462 9561: gap of unknown length
9562 11435: contig of 1874 bp in length
11436 11535: gap of unknown length
11536 13535: contig of 2000 bp in length
13536 13635: gap of unknown length
13636 17233: contig of 3688 bp in length
17234 17424: gap of unknown length
17425 20406: contig of 2963 bp in length
20407 20506: gap of unknown length
20507 23445: contig of 2939 bp in length
23446 23545: gap of unknown length
23546 25600: contig of 2055 bp in length
25601 25700: gap of unknown length
25701 28049: contig of 2349 bp in length
28050 30474: gap of unknown length
30475 30574: contig of 2325 bp in length
30575 35498: gap of unknown length
35499 35598: contig of 4924 bp in length
35599 35988: gap of unknown length
35989 39285: contig of 3666 bp in length
39286 41970: gap of unknown length
41971 42070: contig of 2586 bp in length
42071 45659: gap of unknown length
45660 45759: contig of 3589 bp in length
45760 50332: gap of unknown length
50333 50432: contig of 4573 bp in length
50433 54695: gap of unknown length
54696 54795: contig of 4263 bp in length
54796 58289: gap of unknown length
58290 58390: contig of 3494 bp in length
58391 63599: gap of unknown length
63600 63700: contig of 5210 bp in length
63701 70305: gap of unknown length
70306 70405: contig of 6606 bp in length
70406 76123: gap of unknown length
76124 76223: contig of 5718 bp in length
76224 84961: gap of unknown length
84962 85061: contig of 8738 bp in length
85062 93614: gap of unknown length
93615 93714: contig of 8553 bp in length
93715 103352: gap of unknown length
103353 103452: contig of 9638 bp in length
103453 110399: gap of unknown length
110399 116180: contig of 6847 bp in length
116181 116280: gap of unknown length
116281 125484: contig of 5781 bp in length
125485 125584: gap of unknown length
125584 125584: gap of unknown length

FEATURES	*	125585	138792:	contig of 13208 bp	in length
source			Location/Qualifiers		
		1.	138792		
		"/Organism="Rattus norvegicus"			
		"/db_xref="taxon:10116"			
		"/clone="CH30-320N23"			
BASE COUNT		34391	a	31996	c 31712 g 36193 t 4500
OTHER		otherseq			

Query Match	42.9%	Score 501	DB 2	Length 138792
Best Local Similarity	83.6%	Pred. No. 7e-121		
Matches 628	Conservative	0	Mismatches 115	Indels 8
				Gaps 5

OY	425	CAGGACCACTTTGATGAGGGGAAAGCGTGTACTGAAAGCTGGACATTGGCTGGGAAGGGT	484
Db	137637	CAGGTGCACTTTGATGAGGGGAAAGCGTGTACTGAAAGCTGGACATTGGCTGGGAATGGT	137686
OY	485	GTCGTGGCCCTGGCCCTGCTGCTGGAAGAAATTCTCAGCCACACAGCAAGCTTCTCCGGGGCC	544
Db	137687	GTGGCGGGCCCTGGCCCTGCTGCTGGAAGAAATTCTCAGCCACACAGCAAGCTTCTCCGGGGCC	137746
OY	545	CAGCTCCGTTTGTGCCAGAGTGTCTGGGCTGTGTGCCCTGGCGCAGGGGTCTTCCCTCGG	604
Db	137747	CAGCTCCGTTTGTGCCAGAGTGTCTGGGCTGTGTGCCCTGTGCAGCAGGGGTCTTCCCTCGG	137806
OY	605	ATTCGGACCCCTCCCTGGGGCACTTAAAGCGTGGCCCCCTTCTTAACCTACTTTGGACATC	664
Db	137807	ATTCGGTACCATACCTTTCAGTCATATTAAAGCGGGCTACTCTGTGTGCATGAAATTTGCTC	137866
OY	665	TTTCAAGTTCACTAGAGGGGCTTGTCTCTCCAGATTTCTTAAACTTTTCCC---TGGCTC	720
Db	137867	CACAGAAACAAAGAGAGATCTGTGATGACAGACTTCTATACTCCACATATACAGATAT	137926
OY	721	CAGGAGCATCACACACCTCCCTCAACCCACCCCACTCTGTGCACCCCTCGCTGCTCCTT	780
Db	137927	CCAGACCCACACACACACTCCCCACCCCAAGCCACACTCTCCACCCCTAGCTGTCTCTT	137986
OY	781	GGTCCAGTCCGCT--CTCTCTCTCAAGGCAAGCGAGAGCTGTGCATAG--TTTCCATTTCCA	838
Db	137987	GCTCCAGTCCGCTCTCTCTCTCTTCCCAAGGCAAGCGAGAGCTGTATACATGTTTTCCATTCCA	138046
OY	839	CAGACGATTCCTTCTCTCTCTCTTCT--TAACTATCCATCCCAACCAACTATATCCACTAGTAC	897
Db	138047	CAGACATATTTCCTGCTTCTCTGTATACCTCCACCCCAACCACTAGTGGCCGCTCAGAGAC	138106
OY	898	TCCCAAGCCCCCTACTTATTCCTGACTTCCCCACCCCACTCACCCGACACAGTGTATTATG	957
Db	138107	TCCCAAGCCCCCCCAACGTA--CCCGACTACCCGCCGCCCAACCTCGGAGACCTGTTTTAACTG	138165
OY	958	ACTTTGTCCACAGGACACTGAGATGGGCTGTGACTGTGTGGCAGAGAACCCAGAAACCTGG	1017
Db	138166	ACTATGTGCATCAGGCACTCAAGATTAAGGCTGTGACTGTGTGGCAGAGAACCCAAAGAACTGG	138225
OY	1018	GACTAAGGCACAAAGTCTCCCACTGTGAGGGGAAAGACTGGGGCAAGCTCTCCCTGGGA	1077
Db	138226	GACTAAGGCACAAAGTCTCCCAATGTGAGGGGAAAGAGATGTGAGCAAGCTCTCCCTGGGA	138285
OY	1078	TCCCTGTGATTTTGAAAAGATACTATTTTATTTATTTATTTGTGCACAAATGTTAAATGGA	1137
Db	138286	TCCCTGTGATTTTGAAAAGATACTATTTTATTTATTTATTTGTGCACAAATGTTAAATGGA	138345
OY	1138	TATTTAAGAGAAATTAATCATGATTTCTCTTC	1168
Db	138346	TATTTAAGAGAAATTAATCATGATTTCTCTTC	138376

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS
AX180714	Sequence 1 from patent WO0145730.	AX180714	AX180714.1	GI:15132570

SOURCE	synthetic construct.
ORGANISM	synthetic construct
REFERENCE	artificial sequences.
AUTHORS	1 (bases 1 to 858)
TITLE	Wiley, S.R.
JOURNAL	Twink receptor
	Patent: WO 0145730-A 1 28-JUN-2001
	IMMUNEX CORPORATION (US)

Query Match	42.7%	Score 498.8	DB 6	Length 898
Best Local Similarity	87.0%	Pred. No. 1.4e-120		
Matches 548; Conservative	0	Mismatches 82	Indels 0	Gaps 0

OY	56	AGCTGGGGAGACTGGGGACAAGCCTGTCTGCCAGGAGACCTTCTCAGGAGAGAGCTGACAGCA	115
Db	250	AGTTTTGGGGAGACCCGGGGATCTCTCTCCGCCAGGAGGCTGGCCAGGAGAGCTGTGTGGCA	309
OY	116	GAGGACCCCGGAGAGCCCTCGAAGCTGAATCCCCAGACAGAGGAACCCAGAGATGTGGTA	175
Db	310	GAGGAGAGACAGAGACCCCTCGGAAGTGAATCCCGACAGAGAAAGAACCCAGAGATCTTCGC	369
OY	176	CCTTTCTTGGAAACAATAATGTCGGGCTCGAAGAAATGCTCTCTAAAGCCGGAAAGGCGGG	225
Db	370	CCTTTCCCTGAACCGGACTGTGTGGCCTTCGCAAGAGTGACACTAAAGGCCGGAAACACGG	429
OY	236	CCTGCGGAGGATATTGGACCGCCATTATGAGTTGATCATCTCGGCCACGAGACAGATGGAGCA	295
Db	430	GCTCGAAGAGGATCGCAGCCCATTAATGAACTTATCCAGACCTTGAGCAGAGCGAGCG	489
OY	296	CAAGCAGGTGTGATGGAGGACAGTAGTGCGCTGGGAAAGACAAATAATCAACAGCTCCAGC	355
Db	490	CAGGACAGGTGTGAGACGGGACAAGTAGTGCTGGGAGGAAGCAGAAATCAACAGCTCCAGC	549
OY	356	CCTCTGCGCTACGACCCGCAATTTGGGGAATTTACAGTCAATCAGGGCTGGGCTCTACTAC	415
Db	550	CCTTCGCCCTCAACCCGGCAATGGGGAGTTTATAGCACCCCGGCTGGGCTCTACTAC	609
OY	416	CTGTACTGTCAAGGTCACATTTGATGAGGGAAGGCTGTCTACTCGAAGCTGGAATTTGCTG	475
Db	610	CTGTACTGTCAAGGTCACATTTGATGAGGGAAGGCTGTCTACTCGAAGCTGGAATTTGCTG	665
OY	476	GTGAACGGTGTGCTGGCCCTGCGCTGTGGAAGAATTTCTCAGCCACAGCAGCAAGCTCT	535
Db	670	GTCGATGTGTGTCGTGGCCCTCGCTGCCCTGGAGGAATTTCTCAGCCACAGCAGCAAGCTCT	729
OY	536	CCTGGGCCCCAGACTCCGTTTGTGCCAGGTGTCTGGGCTGTGGCCGTGCGGCCAGGGTCT	595
Db	730	CTCGGGCCCCAGACTCCGCTCTGCCAAGTGTCTGGGCTGTGGCCGTGCGGCCAGGGTCT	789
OY	596	TGCCCTTGGATTCGACACCTCCCTGGGCTATCTTTAAAGCTGCCCCCTTCTAACTAC	655
Db	790	TGCCCTTGGATTCGACACCTCCCTGGGCTATCTTTCAAGGCTGCCCCCTTCTCACTAC	849
OY	656	TTTGGACTCTTTCAAGTTCACTGAGGGGCC	685
Db	850	TTTGGACTCTTCAAGTTCACTGAGGGGCC	879

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan. 1 (bases 1 to 215795)
Akher, N., Antonellis, A., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Latic, P., Lee-Lin, S.-O., Legaspi, R., Maduro, O.L., Maduro, V.B., Maroules, E.H., Masello, C., Maskeri, B., Mastrian, S.D., McInroy, M.E., Prasad, A., Schueler, M.G., Stantir, S., Thomas, J.W., Thomas, P.J., Touchman, J.W., Tsurgou, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 215795)
Green, E.D.
Direct Submission
Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA
Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoo@nhgri.nih.gov
Project Information
Center project name: cms
Center clone name: 145D13
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 203929 bases at least Q40
Consensus quality: 206865 bases at least Q30
Consensus quality: 209181 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 214395; sum-of-ctrls
Quality coverage: 7.30x in Q20 bases; agarose-fp
Quality coverage: 6.47x in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1
3593: contig of 3593 bp in length
3594
3694: gap of unknown length
8001: contig of 4308 bp in length
8002
8101: gap of unknown length
15639: contig of 7538 bp in length
15640
15739: gap of unknown length
15740
24775: contig of 9036 bp in length
24776
24875: gap of unknown length
24876
30150: contig of 5275 bp in length
30151
30250: gap of unknown length
30251
39212: contig of 8962 bp in length
39213
39312: gap of unknown length
39313
49342: contig of 10030 bp in length
49343
49442: gap of unknown length
49443
60956: contig of 11514 bp in length
60957
61056: gap of unknown length
61057
74520: contig of 13464 bp in length
74521
74620: gap of unknown length
74621
89522: contig of 14902 bp in length
89523
89622: gap of unknown length
89623
105537: contig of 15915 bp in length
105538
105637: gap of unknown length
105638
121197: contig of 15560 bp in length
121297: gap of unknown length
121298
138883: contig of 17586 bp in length

FEATURES
source
1. 215795
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-145D13"
/clone="RP43"
1. 3593
/note="assembly-fragment"
3694. 8001
/note="assembly-fragment"
8102. 15639
/note="assembly-fragment"
15740. 24775
/note="assembly-fragment"
24876. 30150
/note="assembly-fragment"
30251. 39212
/note="assembly-fragment"
39313. 49342
/note="assembly-fragment"
49443. 60956
/note="assembly-fragment"
61057. 74520
/note="assembly-fragment"
74621. 89522
/note="assembly-fragment"
89623. 105537
/note="assembly-fragment"
105638. 121197
/note="assembly-fragment"
121298. 138883
/note="assembly-fragment"
138984. 166802
/note="assembly-fragment"
166903. 215795
/note="assembly-fragment"

BASE COUNT 52160 a 54859 c 53752 g 55513 t 1511 others
ORIGIN

Query Match 26.0%; Score 304; DB 2; Length 215795;
Best Local Similarity 71.7%; Pred. No. 7.9e-69;
Matches 501; Conservative 0; Mismatches 155; Indels 74; Gaps 11;

QY 425 CAGTGCACCTTGATGAGGAAAGGCTGTCTACCTGAGCTGAGTGTGGAAGGT 484
|||||
DB 199456 CAGTGCACCTTGATGAGGAAAGGCTGTCTACCTGAGCTGAGTGTGGAAGGT 199515
QY 485 GTGCTGCGCCCTGCGCTGCTGGAAGAATCTACGACAGACAGAACTCTCGGGCCC 544
|||||
DB 199516 GTGCTGCGCCCTGCGCTGCTGGAAGAATCTACGACAGACAGAACTCTCGGGCCC 199575
QY 545 CAGCTCGTTTGGCCAGGAGTCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 604
|||||
DB 199576 CAGCTCGTTTGGCCAGGAGTCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGG 199635
QY 605 ATCCGACCCCTCCCTGCGCTATCTTAAGGCTGCCCCCTTCTTAACCTACTTTGGACTC 664
|||||
DB 199636 ATCCGACCCCTCCCTGCGCTATCTTAAGGCTGCCCCCTTCTTAACCTACTTTGGACTC 199695
QY 665 TTTCAGTCACTGAGGAGCTTGTCTCTCCAGATTCCTTAAACTTTCCCTGGCTCCAGG 724
|||||
DB 199696 TTTCAGTCACTGAGGAGCTTGTCTCTCCAGATTCCTTAAACTTTCCCTGGCTCCAGG 199752
QY 725 AGATCAACCAACCTCCACACCCACCCACCTCCACCCACCCCTC-GGTGGCTTGGT 783
|||||
DB 199753 --CTGACAGCTCTTGGGACCCGCTCTCTGCCCCACCCCTTACGCTGCTTTGGCT 199810
QY 784 CCAGTCTGTCTCTCC--TCAAGGACGACAGGCTTTTCAATG-TTTCATTCACAA 840
|||||

Db	199811	CCAGACACGTGCCCCCTGCCCTCTGTAGAGGCTGCGTGGGCTGTTCACGTGTTCATCCACACA	199870
Qy	841	GACGTA-----TCCTTGCCTCTCTTAACATCCATCCACACACACTATACCTGCAC	893
Db	199871	TAAATACAGTATATCCCATCTTATTTACAACTCCCCACCGCCACCTCTCCACTGCAC	199930
Qy	894	TAGCTCCCAAGGCCCTAC-----TTATCCCTGACTCCCCACCCACATCA	938
Db	199931	TAGCTCCCCAATCCCTGCACCCCTTTGAGGCCCCCCAGTAGATCTGCACCTCCCCCTGGCCACA	199990
Qy	939	CCCGACACAGTGTATTATGACTTTGTGCAC-----	968
Db	199991	GACCCCCAGGGGATGTCTTACACTGTACTGTGGGCAAGATGGGTCCAGAAAGCCCA	200050
Qy	969	---CAGGCACGTAGATGGGCTGACCTGTGGCAGAGAACCCAGAGACCTGGGCTAGGC	1025
Db	200051	CTTCAGGCACTAAGAGGGGCTGTGACCTGGCGCAGAGAACCCAAAGATGCTGGGCTTAGGC	200110
Qy	1026	CAGAGATCCCAACGTGTAGGGGGGAAGAGCTGGGACAAAGCTCCGCCGTGA---TCGC	1081
Db	200111	CAGGAGTTCGCCAATGTGTAGGGGGCAGAAACACA--GACACAGCTCTCTCCCTTGGAAATTTCC	200169
Qy	1082	TGTGATTTTGA---AGATACTATTTTATTTATTTGTGACAAATGT---TAAATGG	1136
Db	200170	TGTGATTTTAAAACAGATATTTATTTATTTATTTGTGACAAAGTGTGATTAATGG	200229
Qy	1137	ATATTAAAGAGATTAATCATGATTTCTCT	1166
Db	200230	ATATTAAATAGATATAGTCATAGCTCTCTCT	200259

RESULT 12	AC016876/c	LOCUS	AC016876	177703 bp	DNA	1 linear	HTG	06-AUG-2002
DEFINITION			Homo sapiens clone RP11-186B7, ***	SEQUENCING IN PROGRESS ***	10			
			unordered pieces.					

```

VERSION      AC016876.5  GI:2131830
KEYWORDS     HTG; HTGS_PHASE1; HTGS_FULLTOP; HTGS_ACTIVEFIN
SOURCE       human.
ORGANISM     Homo sapiens

```

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS

1 (bases 1 to 177703)
Bliren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome, clone RP11.16687
unpublished
2 (bases 1 to 177703)
Bliren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,

TITLE Direct Submission
JOURNAL Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE 3 (bases 1 to 17703)
AUTHORS Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,

Barna, N., Baastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarate, J., Chang, J., Chazaro, B., Choepel, Y., Collimore, A., Cook, A., Cooke, P., DeArrelano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Garry, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gird, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Morton, D., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,

Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mhova, T., Mlewa, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Plunkharg, P., Plerre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupp, B., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, B., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemek, L., Zimmer, A. and Zody, M.

Direct Submission

TITLE
JOURNAL
COMMENT

Submitted (06-MGC-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 2, 2002 this sequence version replaced g1:35421989.

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIGR
 Web site: <http://www.seq.wi.mit.edu>
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: I3849
 Center clone name: 186_B7

```

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number WILL
* be preserved.

```

*	1	48645:	contly of 48645	bp	in length
*	4866	48745:	gap of	100	bp
*	48746	58227:	contly of 9552	bp	in length
*	58278	58377:	gap of	100	bp
*	58378	65804:	contly of 7427	bp	in length
*	65805	65904:	gap of	100	bp
*	65905	79793:	contly of 13889	bp	in length
*	79794	79893:	gap of	100	bp
*	79894	99493:	contly of 19600	bp	in length
*	99494	99593:	gap of	100	bp
*	99594	11104:	contly of 11456	bp	in length
*	111050	111149:	gap of	100	bp
*	111150	125020:	contly of 13871	bp	in length
*	125021	125120:	gap of	100	bp
*	125121	145109:	contly of 19989	bp	in length
*	145110	145209:	gap of	100	bp
*	145210	169456:	contly of 24249	bp	in length
*	169459	169558:	gap of	100	bp
*	169559	177703:	contly of 8145	bp	in length

```

FEATURES
  source      Location/Qualifiers
1. 177703    /organism="Homo sapiens"
          /db_xref="taxon:9606"
          /clone="RP11-186B7"
          /clone_11b="RPIC1-11 Human Male BAC"
BASE COUNT  42862 a 45967 c 46586 g 41254 t 1052 others
ORIGIN

```

	Query Match	Similarity	26.0%	Score 303.8	DB 2:	Length 177703
	Best Local	Similarity 71.0%	Pred: No.	8.7e-69		
	Matches 574	Conservative	0	MissMatches 162	Indels 73	Gaps 10
Oy	425	CAGGTGCACATTGATGAGGGAAGCGTCGTCTACCTGAACTGGACTTGCCTGTAACGCT	484			
Db	37314	CAGGTGCACATTGATGAGGGAAGCGTCGTCTACCTGAACTGGACTTGCCTGTAACGCT	37255			
Oy	485	GTGCTTGCCCTGCGCTGCGCTGGAAGAATTCTCAGGCACACGACGAAGCTCTCTCTGGGCC	544			

```

Db 37234 GTGTCGCCCTGCGCTGCGCTGGAGGAATTCACAGCAGCTGCGGCGAGTTCCCTCGGGCC 37195
Oy 545 CAGCTCCGTTGTGTCAGAGTGTCTGGGGCTTGCCGCTGCGGCGAGGGCTTCCTTCGG 604
Db 37194 CAGCTCCGCTCTGCGAGAGTGTCTGGGGCTTGCCGCTGCGGCGAGGGCTTCCTTCGG 37135
Oy 605 ATCCGACCCCTCCCTGGGCTTCATCTTAAGGCTGCCCCCTTCCTTAACCTACTTCTTGAGTC 664
Db 37134 ATCCGACCCCTCCCTGGGCGCATCTCAAGGCTGCCCCCTTCCTTAACCTACTTGGAGTC 37075
Oy 665 TTTCAGTTCACAGAGGGGCTTCTCTCCAGATTCCTTAACCTTCTGCTGCTCAG 724
Db 37074 TTTCAGTTCACAGAGGGGCTTCTCTCCAGATTCCTTAACCTTCTGCTGCTCAG 37018
Oy 725 AGCATCACACACCTCCCTACCCCGACCTCTCCACCCCGCTC-GCTGCTCTGCT 783
Db 37017 --CCTCGACAGCTCTCTGCGGCGCGGCTCCCTGCGGCGCATCTCAGCGCTCTTGT 36960
Oy 784 CCAGTCTCTCTCTCTC--TCAAGGCGACGAGAGCTTGTTCACATGTTTCATTC-- 837
Db 36959 CCAGACCTGCGCCCTCCCTCTAGAGGCTGCTGGGCTGTACGTTTTCATCCACACA 36900
Oy 838 ---ACAGACGATCTCTCTCTCTTAAACATCCCATCCACACACTATCCACTACT 894
Db 36899 TAAATACAGTATTCCTCTATCTTACACTCCCGACCGCCCATCTCTCAGCTCACT 36840
Oy 895 AGCTCCCAAGGCCCTAC-----TTATCCCTGACCTCCCGACCCAGCTCAG 939
Db 36839 AGCTCCCAATTCCTGAGCCCTTTGAGGCCCGCCAGTGAATCTGACCTCCCGCTGCGCAG 36780
Oy 940 CCGACACAGCTGTTTATGACTTTTGTCAC----- 968
Db 36779 ACCCCAGGAGCATTTGTTACACTGTACTGTGGGCAAGAGTGGTCCAGAACCCAC 36720
Oy 969 --CAGGACTGATGAGCTGAGCTGTGTCAGAGACGAGACCTGGGCTAGGCC 1026
Db 36719 TTCAGGCACTAAGAGGGCTGACCTGGCGGAGAGACCAAGAGACTGGGCTAGGCC 36660
Oy 1027 AGAAGTTCACACTGTGAGGGGAGAGAGCTGGGCAAGCTCTCCCTGGA----TCCCT 1082
Db 36659 AGGAGTTCACAAATGTGAGGGGCGAGA--AACAGACAAAGCTCTCCCTGAGAAATTCCT 36601
Oy 1083 GTGATTTTGA--AGATACTATTTTATTTATTTATTTGACAAATGT---TAAATGA 1137
Db 36600 GTGATTTTAAACAGATATATTTTATTTATTTATTTGACAAATGTGATTAATGA 36541
Oy 1138 TATTAAGAGATTAATCATGATTTTCT 1166
Db 36540 TATTAATAGATATAGTATGATCTCT 36512

RESULT 13
AC126921 153553 bp DNA linear HTG 10-JUL-2002
LOCUS AC126921
DEFINITION Bos taurus clone RP42-45D24, WORKING DRAFT SEQUENCE, 12 unordered
pieces.
ACCESSION AC126921.1 GI:21724098
VERSION HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 153553)
REFERENCE
AUTHORS
Akter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-I., Idol,J.R., Karlins,E., Lalic,P.,
Lee-Jin,S.-O., Legaspi,R., Maduro,Q.L., Maduro,V.B.,
Maguiles,E.H., Mastello,C., Maskeri,B., Mastrian,S.D.,
McCluskey,J.C., McDowell,J., Paguirigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stantirip,S., Thomas,J.W.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT
Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A.,
Welchby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 153553)
Green,E.D.
Direct Submission
Submitted (10-JUL-2002) NIH Intramural Sequencing Center, 8717
Groveomont Circle, Gaithersburg, MD 20877, USA
-----
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoehqri.nih.gov
-----
Project Information
Center project name: ddi
Center clone name: 045D24
-----
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 14606 bases at least Q40
Consensus quality: 147748 bases at least Q30
Consensus quality: 148824 bases at least Q20
Insert size: 151000; agarose-fp
Insert size: 152453; sum-of-ctnigs
Quality coverage: 8.80x in Q20 bases; agarose-fp
Quality coverage: 8.72x in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2313 2312: contig of 2312 bp in length
2413 2412: gap of unknown length
2413 5841: contig of 3429 bp in length
5842 5941: gap of unknown length
5942 8435: contig of 2494 bp in length
8436 8535: gap of unknown length
8536 15799: contig of 7264 bp in length
15800 15899: gap of unknown length
15900 25224: contig of 9325 bp in length
25225 25325: gap of unknown length
25325 32505: contig of 7180 bp in length
32505 32604: gap of unknown length
32605 40970: contig of 8366 bp in length
40971 41071: gap of unknown length
41071 56590: contig of 15520 bp in length
56591 56690: gap of unknown length
56691 73769: contig of 17079 bp in length
73770 73869: gap of unknown length
73870 90859: contig of 16990 bp in length
90860 90959: gap of unknown length
111428 111428: contig of 20469 bp in length
111429 111528: gap of unknown length
111529 111529: contig of 42025 bp in length.
Location/Qualifiers
1. 153553
/organism="Bos taurus"
/db_xref="taxon:9913"
/clone="RP42-45D24"
/clone_id="RP42"
1. 2312
/note="assembly-fragment"
2413. 5841
/note="assembly-fragment"
clone_end=77
vector_side="left"
5942. 8435
misc_feature
misc_feature
misc_feature
misc_feature

```


*	21302	21401	gap of unknown length
*	21402	33719	contig of 12318 bp in length
*	33720	33819	gap of unknown length
*	33820	45054	contig of 11235 bp in length
*	45055	45154	gap of unknown length
*	45155	58724	contig of 13570 bp in length
*	58725	58824	gap of unknown length
*	58825	76670	contig of 17846 bp in length
*	76671	76770	gap of unknown length
*	76771	115165	contig of 38395 bp in length
*	115166	115265	gap of unknown length
*	115266	177555	contig of 62290 bp in length

```

/misc_feature 1..3460
/misc_feature 3561..6869
/misc_feature 6970..13465
/misc_feature 13566..21301
/misc_feature 21402..33719
/misc_feature 33820..44504
/misc_feature 44515..58724
/misc_feature 58825..76670
/misc_feature 76771..115165
/misc_feature 115266..177535
/organism="Sus scrofa"
/db_xref="taxon:9823"
/clone="RP4-436K21"
/clone_lib="RP4"

```

[illegible]

Query Match	20.3%;	Score 237.4;	DB 2;	Length 177555;
Best Local Similarity	63.8%;	Pred. No. 3e-51;		
Matches 487; Conservative	0;	Mismatches 216;	Indels 60;	Gaps 6

OY	415	CCCTACTGTCAGGTGCACATTGATGAGGAAAGCGTCGTACTGAGCTGACATTGCT	474
Db	114241	CCGCCCCCCCCAGTGCACCTTTGATGAGGGAAGGCCGCTTACTGAGCTGGAATTGCT	114182
OY	475	GGTGAACGATGTCGTGGCCCTGGCGCTCCGGAAGAAATTCATGACCCACACAGCAAGCTC	534
Db	114181	GGTGAACGACACGCTGGCCCTGGCGCTCCGGAAGAAATTCCTGGCCACGCGCGAGATTTC	114122
OY	535	TTCGGGCCCCAGTCCGCTTTGTCGACAGGTGTCGGGCTTCTCCGCTGGCGGCGAGGCTC	594
Db	114121	CTTGGGCCCCAGTCCGCTCTGTCGACAGGTGTCGGGCTTCTCCGCTGGCGGCGAGGCTC	114062
OY	595	TTCCTTCGAGTCGACACCCCTCCCTGAGGTCACTTTAAGCTGCCCCCTTCCTAACTGA	654
Db	114061	CTCCCTCGGATCGCACCCCTCCCTGAGGCCACACTCAAGGTGCCCCCTTCCTCAACTGA	114002
OY	655	CTTGGGACCTCTTCAATTTACATGAGGGGCGCTTGCTCTCCACAGATTCTTAAACTTTCC	714
Db	114001	CTTGGGACCTCTTCAATTTACATGAGGGGCGCGGGCTTC-----	113962

Oy 715 TGCGTCCAGGAGCATCACACACTTCCTACCCCACCACCCCACTCCTCCACCCGCCTGGTG 774
 |||| | ||| | ||| | ||| | ||| | ||| |
Db 113961 -----CCAGCAGCGGCCCCACCCCGGGGTGCTCCTCCCCAGACCTGCTCCTCCCGTG 113907

OY	775	CTCCCTGGTCCAGTCGTCCTGTCTCTCTCAAAAGGCACGCCAGACTGTTTGACATGTTTCCAT	834
Db	113906	TAGCGTTCCCGGGGTGGGGTCGCTCAGCATGATGTACCCACAGAGATCTCCATGGGTCTCG	113847
OY	835	TCCACAGACGTAATCCCTTCTCTCTTAACATATCCATCCATCCACCACACAATATCCACCTCACT	894
Db	113846	ACAAAC-----TCCCCCACCTCTGGCTCTCCCCCTCACT	113812
OY	895	AGCTCCCAAAGCCCCTACTTAATCCCTGCATCCCCACCACCACTCACCCGACACCACTGTTTTA	954
Db	113811	GGCTTCACAAAACCCCATTTAGCTCTCTGGCCGACGCCACAGGACACTGCACTCTGGGTGA	113752
OY	955	TTGACTTTGTGCACACAGGACACTGAGATGGGCTGGACCTGTGTGCAGAAAGCCAG-AGAAC	1013
Db	113751	CATGGGTCCCAACACCGCTGTTCAGGCACTAAGAGAAAGGCGGAGGCCACAGAGAGCCCGGG	113692
OY	1014	CTGGGACTAGGCGACGAAGATTCTCCCAACTGTGAGGGGGAAGAGCTGGGAGCAACGCTCTCCC	1073
Db	113691	CCGGGAAGAGGCTGGGAAATTTCGAAATTTTCAGAGGGAAGAGCCCAAGACGAGCCCTCCC	113632
OY	1074	TGGA----TCCCTGTGATTTTTGAAA--AGATACTATTTTTATTAATTAATGTGACAAAT	1127
Db	113631	CTGATCAATCCCTGTGGATTTTTAAACAGATATATTTTTTAATTAATGTGACAAAT	113572
OY	1128	GTT---AAATGATATTAAAGAGAAATAATGATATTTCTCTT	1167
Db	113571	GTTGAGAGGTGGACATTAAAGAGATTAAAGCCAGGCCGCTCTT	113529

RESULT	15			
AC126925/c				
LOCUS	AC126925	161428 bp	DNA	linear
DEFINITION	Canis familiaris clone RP81-332E11, WORKING DRAFT SEQUENCE, 15			
ACCESSION	AC126925			
VERSION	AC126925.1	GI:21724102		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.			
SOURCE	dog.			
ORGANISM	Canis familiaris			

REFERENCE	TITLE	REFERENCE	JOURNAL	COMMENT
1 (bases 1 to 161428)				
Ahtler, N., Antonellis, A., Ayele, K., Beckstrom-Stenberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G., Breon, K., Brinkley, C., Brooks, S., Dietrich, N.U., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Idol, J.R., Karlins, E., Laric, P., Lee, L.N., S.-Q., Legaspi, R., Maduro, Q. L., Maduro, V. B., Margulies, E. H., Mastello, C., Maskeri, B., Mestrarian, S. D., McCloskey, J. C., McDowell, J., Padirigan, C., Pearson, R., Portnoy, M. E., Prasad, A., Schneller, M. G., Stantford, S., Thomas, J. W., Thomas, P. J., Touchman, J. W., Tsuregon, C., Vogt, J. L., Walker, M. A., Wehedy, K. D., Wiggins, L., Young, A., Zhang, L. H. and Green, E. D.				
NISC Comparative Sequencing Initiative				
Unpublished				
2 (bases 1 to 161428)				
Green, E. D.				
Direct Submission				
Submitted (10-JUL-2002)				
NTH Intramural Sequencing Center, 8717				
Grovemont Circle, Gaithersburg, MD 20877, USA				
----- Genome Center				

```

Center ode: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
-----
Project Information
Center project name: cwp
Center clone name: 332E11
-----
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: dye-terminator big dye; 100% of reads
Assembly program: phrap; version 0.990319
Consensus quality: 148712 bases at least Q40
Consensus quality: 151471 bases at least Q30

```


GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 11:50:36 ; Search time 66 Seconds
(without alignments)
573.381 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 284
Sequence: 1 MSLDFEISARRLPLPSISG.....PMAHLKAAPFLYFGFLQVH 284

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 65 summaries

Database :

A_Geneseq_101002.*
1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	284	100.0	284	19 AAM47525	Homo sapiens tumou
2	249	87.7	249	20 AAB09369	Human tumour necro
3	249	87.7	249	21 AAB07526	Amino acid sequenc
4	249	87.7	249	21 AAY95338	Human PRO207 antit
5	249	87.7	249	23 AAB06129	Human PRO207 polyp
6	241	84.9	249	19 AAM29745	TNF related endoth
7	241	84.9	249	19 AAB00891	Human TREPA (TNF r
8	207	72.9	273	22 AAB03439	TWEAK extracellular
9	146	51.4	146	22 AAB00895	Human TREPA (TNF r
10	143	50.4	189	19 AAM29746	TNF related endoth

ALIGNMENTS

RESULT 1	ID	AAW47525	standard; Protein: 284 AA.
XX	XX	AAW47525:	
AC	XX	21-JUL-1998 (first entry)	
DT	XX		
XX	XX		
DE	XX	Homo sapiens tumour necrosis factor related ligand (TRELL).	
XX	XX	TRELL: tumour necrosis factor related ligand; tnfr; treatment;	
KW	XX	cancer; autoimmune disease; immune system; stimulation; suppression;	
KW	XX	graft rejection.	

11	143	50.4	189	22	AAB00892	Human UL4flap TRP
12	107	37.7	208	20	AAM3590	Human TNRL3 protei
13	46	16.2	211	20	AAM3591	Mouse TNRL3 protei
14	32	11.3	225	19	AAM47524	Mus musculus tumou
15	32	11.3	225	19	AAM47527	Amino acid sequenc
16	9	3.2	58	21	AAG01265	Human secreted pro
17	9	3.2	748	20	AAI14906	Human secreted pro
18	9	3.2	748	20	AAI14906	Extended amino aci
19	9	3.2	749	23	AAB73512	M vaccae GYS-9 pro
20	8	2.8	749	23	AAB73512	Protonibacterium
21	8	2.8	55	22	AAB42637	Peptide #10143 enc
22	8	2.8	55	22	AAB25983	Protein #7982 enco
23	8	2.8	55	22	AAM63528	Human brain expres
24	8	2.8	55	22	AAM76342	Human bone marrow
25	8	2.8	55	22	AAM20858	Peptide #7292 enco
26	8	2.8	55	22	AAM36451	Peptide #10488 enc
27	8	2.8	55	22	AAB45635	Human peptide enco
28	8	2.8	65	22	AAB43981	Peptide #11487 enc
29	8	2.8	65	22	AAB26890	Protein #8889 enco
30	8	2.8	65	22	AAM64984	Human brain expres
31	8	2.8	65	22	AAM7706	Human bone marrow
32	8	2.8	65	22	AAM7706	Peptide #8055 enco
33	8	2.8	65	22	AAM37923	Peptide #11960 enc
34	8	2.8	65	22	AAB46737	Human peptide enco
35	8	2.8	69	22	AAB03723	Novel human diagno
36	8	2.8	71	22	AAB03663	Novel human diagno
37	8	2.8	84	22	AAB01282	Protonibacterium
38	8	2.8	110	22	AAB02059	Novel human diagno
39	8	2.8	117	22	AAB08094	Human polypeptide
40	8	2.8	184	22	AAB67712	Drosophila melanog
41	8	2.8	190	23	AAB28041	Streptococcus poly
42	8	2.8	198	22	AAM23684	Human EST encoded
43	8	2.8	222	23	ABG1674	Human ovarian anti
44	8	2.8	266	22	ABG16279	Novel human diagno
45	8	2.8	307	22	ABG01186	Novel human diagno
46	8	2.8	370	22	ABG05012	Novel human diagno
47	8	2.8	370	22	ABG13391	Novel human diagno
48	8	2.8	370	22	ABG18115	Novel human diagno
49	8	2.8	424	22	ABG15613	Novel human diagno
50	8	2.8	430	22	AAM37552	Human EST encoded
51	8	2.8	454	22	AAB57908	Drosophila melanog
52	8	2.8	586	22	ABG20260	Novel human diagno
53	8	2.8	592	22	AAU32148	Novel human secret
54	8	2.8	603	22	ABG03722	Novel human diagno
55	8	2.8	633	22	ABG20261	Novel human diagno
56	8	2.8	799	22	ABG28291	Novel human diagno
57	7	2.5	12	16	AAB66780	Cell adhesion prot
58	7	2.5	17	23	AAB09088	Human lymphotoxin
59	7	2.5	18	23	ABG62026	Human tumour suppr
60	7	2.5	23	23	ABG62030	Human tumour suppr
61	7	2.5	24	20	AAV4189	Human pre-prohepar
62	7	2.5	25	21	AAB18622	Antigenic peptide
63	7	2.5	44	20	AAB89072	Polyptide fragment
64	7	2.5	44	22	AAB51243	Human secreted pro
65	7	2.5	44	22	AAB01449	Human gene 14 enco

AC	AAB07526;
XX	
DT	20-OCT-2000 (first entry)
XX	
DE	Amino acid sequence of a soluble recombinant human TWEAK protein.
XX	
XX	TWEAK protein; immunological disorder; immune response; inflammation;
KW	TWEAK blocking agent; autoimmune disease; organ transplant rejection;
KM	Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour.
OS	Homo sapiens.
XX	
PM	WO200042073-A1.
PD	20-JUL-2000.
XX	
PF	14-JAN-2000; 2000MO-US01044.
XX	
PR	15-JAN-1999; 99US-0116168.
PA	(BIOJ) BIOGEN INC.
PI	Rennett P;
XX	
DR	WPI: 2000-476036/41.
XX	
PT	Preventing and treating immune responses using modulators, especially
PT	antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for
PT	treating e.g. inflammation and graft versus host disease -
XX	
PS	Disclosure: Flg 1; 45pp: English.
XX	
CC	The present sequence represents a TWEAK protein. The specification
CC	describes a method for preventing or treating an immunological
CC	disorder and/or inhibiting an immune response in an animal. The
CC	method comprises administering a TWEAK blocking agent. The method may
CC	be used for preventing and treating immune disorders associated with
CC	inappropriate expression and/or activity of TWEAK. These disorders
CC	include autoimmune diseases, acute and chronic inflammation, organ
CC	transplant rejection, Graft-versus-host disease (GVHD), lymphoid cell
CC	malignancies, septic and other forms of shock, loss of immune
CC	responsiveness (as seen in human immunodeficiency virus (HIV)
CC	infections) and failure of the immune response to tumour growth.
SQ	Sequence 249 AA;
Query Match	87.7%; Score 249; DB 21; Length 249;
Best Local Similarity	100.0%; Pred. No. 2/e-220;
Matches 249; Conservative	0; Mismatches 0; Indels 0; Gaps 0.
OY	36 MAARSORRRGRGEGTALLVPLALGIGLALACLGILLAVSIGSRASLSAOEPAOEEL 95
DB	1 MAARRSORRRGRGEGTALLVPLALGIGLALACLGILLAVSVLSGRASLSAOEPAOEEL 60
OY	96 VAEEDODPSELNPQTEESODPAPFLNRLVLRPRSAPKGRKTARRAIAAHYEVPHPGOD 155
DB	61 VAEEDODPSELNPQTEESODPAPFLNRLVLRPRSAPKGRKTARRAIAAHYEVPHPGOD 120
OY	156 GAQAGVDGTSGSWEERAINSSSPRLRYNRQIGEFITYTRAGLYTYLCOVHFDESKAYVLKLD 215
DB	121 GAQAGVDGTSGSWEERAINSSSPRLRYNRQIGEFITYTRAGLYTYLCOVHFDESKAYVLKLD 180
OY	216 LLYDVLVLRCLFEFSATMAASSLGQLRLCOVSGLLIAPGSSLRRTLPMAHLTKAAPFL 275
DB	181 LLYDVLVLRCLFEFSATMAASSLGQLRLCOVSGLLIAPGSSLRRTLPMAHLTKAAPFL 240
OY	276 TYFGLEFOYH 284
DB	241 TYFGLEFOYH 249

ID	AA95338	standard; Protein; 249 AA.
XX		
AC	AA95338;	
XX		
DT	25-SEP-2000	(first entry)
XX		
DE	Human PRO207 antitumour protein.	
XX		
KW	PRO207; human; antitumour; tumour; therapy; cytostatic;	
KW	breast cancer; ovarian cancer; renal cancer; colorectal cancer;	
KW	uterine cancer; prostate cancer; lung cancer; bladder cancer;	
KW	central nervous system cancer; melanoma; leukaemia; neoplasm.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FH	Peptide	1..40
FT		/label= Signal_peptide
FT	Protein	41..249
FT		/label= PRO207
FT	Modified-site	27..33
FT		/note= "N-myristoylation"
FT	Modified-site	29..35
FT		/note= "N-myristoylation"
FT	Modified-site	36..42
FT		/note= "N-myristoylation"
FT	Modified-site	45..51
FT		/note= "N-myristoylation"
FT	Modified-site	118..124
FT		/note= "N-myristoylation"
FT	Modified-site	121..127
FT		/note= "N-myristoylation"
FT	Modified-site	125..131
FT		/note= "N-myristoylation"
FT	Modified-site	128..134
FT		/note= "N-myristoylation"
FT	Modified-site	139..143
FT		/note= "Asn is N-glycosylated"
FT	Modified-site	10..14
FT		/note= "amidation"
FT	Modified-site	97..101
FT		/note= "amidation"
FT	Peptide	24..35
FT		/note= "prokaryotic membrane lipidprotein lipid"
PN	WO200037638-A2.	
XX		
PD	29-JUN-2000.	
XX		
PF	02-DEC-1999;	99WO-US28565.
XX		
PR	22-DEC-1999;	98US-0113286.
PR	08-MAR-1999;	99WO-US05028.
PR	21-APR-1999;	99US-0130232.
PR	28-APR-1999;	99US-0131445.
PR	14-MAY-1999;	99US-0134287.
PR	20-JUL-1999;	99US-0144758.
PR	26-JUL-1999;	99US-0145698.
PR	15-SEP-1999;	99WO-US21090.
PR	15-SEP-1999;	99WO-US21547.
XX		
PA	(GETH) GENENTECH INC.	
PI	Ashtenazi AJ, Goddard A, Godowski PJ, Gurney AL, Maisters SA,	
PI	Napier MA, Pitti RM, Wood WI;	
XX		
DR	WPI: 2000-442668/38.	
DR	N-PSDB: AAA49717.	
XX		
FT	Novel composition to inhibit neoplastic cell growth or for treating	
FT	tumor in mammal comprises polypeptides PRO179, PRO207, PRO320, PRO319,	
FT	PRO321, PRO324, PRO328, PRO301, PRO526, PRO362, PRO356, PRO509 or	
FT	PRO866	

XX Claim 19; Fig 4; 172pp; English.
PS
XX
CC The present sequence is that of human antitumour protein PRO207,
CC as deduced from a foetal kidney cDNA clone (see AAA49117). PRO207
CC shows amino acid sequence identity to tumour necrosis factor family
CC members, especially human lymphotoxin-beta (23.4%) and human CD40
CC ligand (19.8%). Mol.wt. is 27,216. A claimed method for inhibiting
CC the growth of a tumour cell comprises exposing the tumor cell
CC to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224, PRO328, PRO301,
CC PRO326, PRO356, PRO509 or PRO866 (see AA195337-49), their
CC agonists or chimeric polypeptides incorporating them. The tumour
CC is especially a cancer selected from breast, ovarian, renal,
CC colorectal, uterine, prostate, lung, bladder and central nervous
CC system cancer, melanoma and leukaemia. Methods for the recombinant
CC expression of the antitumour proteins are also provided.
CC
SQ Sequence 249 AA:

Query Match 87.7%; Score 249; DB 21; Length 249;
Best Local Similarity 100.0%; Pred. No. 2,7e-220;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARSRQRRGRGEGPTALLVPLALGCLALACGLLAAYVSLGSRASLSAQEPAQEEL 95
DB 1 MAARSRQRRGRGEGPTALLVPLALGCLALACGLLAAYVSLGSRASLSAQEPAQEEL 60
QY 96 VAEEDDPSSELNPQTEESODPAFLNRLVPRRSAPKGRKTRARRAIAHYEVHRRPGD 155
DB 61 VAEEDDPSSELNPQTEESODPAFLNRLVPRRSAPKGRKTRARRAIAHYEVHRRPGD 120
QY 156 GAAGVDTGTVSGMEARINSSSPFLRYNRQIGEFYTRAGLYLYLCOVHDEKAVYLLKD 215
DB 121 GAAGVDTGTVSGMEARINSSSPFLRYNRQIGEFYTRAGLYLYLCOVHDEKAVYLLKD 180
QY 216 LVDGVLALRCLEFSATASSIGPOLRLCOVSGLLALRPSSLRIRTLPMANHLKAAPFL 275
DB 181 LVDGVLALRCLEFSATASSIGPOLRLCOVSGLLALRPSSLRIRTLPMANHLKAAPFL 240
QY 276 TYFGLFQVH 284
DB 241 TYFGLFQVH 249

RESULT 5
AAU86129
ID AAU86129 standard; Protein; 249 AA.
XX
AC AAU86129;
XX
DT 15-JUL-2002 (first entry)
XX
DE Human PRO207 polypeptide.
XX
KW Human; PRO; benign tumour; malignant tumour; lymphoid malignancy;
KW leukaemia; neuronal disorder; stromal disorder; blastococelic disorder;
KW inflammatory disorder; immune disorder; anglogenetic disorder;
KW cytoslastic; neuroprotective.
OS Homo sapiens.
XX
PN WO200153486-A1.
XX
PD 26-JUL-2001.
XX
PF 11-FEB-2000; 2000WO-US03565.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 11-MAR-1999; 99US-123972P.
PR 11-MAY-1999; 99US-133459P.
PR 02-JUN-1999; 99WO-US12252.
PR 22-JUN-1999; 99US-140650P.
PR 22-JUN-1999; 99US-140653P.

PR 20-JUL-1999; 99US-144758P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 17-AUG-1999; 99US-149395P.
PR 31-AUG-1999; 99US-151689P.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 05-JAN-2000; 2000WO-US00219.
XX
PA (GETH) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Martens SA, Pan J, Pittl RM, Roy MA, Smith V, Stone DM,
PI Watanabe CK, Wood WI;
XX
PI WPI, 2002-205567/26.
DR N-PSDB; ABR40255.
XX
PT Thirty five nucleic acids encoding PRO polypeptides, useful for
PT treating benign or malignant tumours, leukaemias and lymphoid
PT malignancies, inflammatory, angiogenic and immunologic disorders -
PS Claim 61; Fig 4; 302pp; English.
XX
XX The present invention relates to the isolation of novel human PRO
XX polypeptides and the polynucleotide sequences encoding them. The
XX PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
XX useful for treating benign or malignant tumours (e.g. renal, kidney,
XX bladder, breast, etc), leukaemias and lymphoid malignancies, other
XX disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
XX macrophagal, stromal and blastococelic disorders, inflammatory, immune
XX and anglogenetic disorders. The polynucleotide sequences are also
XX useful in gene therapy. AU86128-AU86162 represent the human PRO
XX polypeptides of the invention.
SQ Sequence 249 AA:

Query Match 87.7%; Score 249; DB 23; Length 249;
Best Local Similarity 100.0%; Pred. No. 2,7e-220;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 MAARSRQRRGRGEGPTALLVPLALGCLALACGLLAAYVSLGSRASLSAQEPAQEEL 95
DB 1 MAARSRQRRGRGEGPTALLVPLALGCLALACGLLAAYVSLGSRASLSAQEPAQEEL 60
QY 96 VAEEDDPSSELNPQTEESODPAFLNRLVPRRSAPKGRKTRARRAIAHYEVHRRPGD 155
DB 61 VAEEDDPSSELNPQTEESODPAFLNRLVPRRSAPKGRKTRARRAIAHYEVHRRPGD 120
QY 156 GAAGVDTGTVSGMEARINSSSPFLRYNRQIGEFYTRAGLYLYLCOVHDEKAVYLLKD 215
DB 121 GAAGVDTGTVSGMEARINSSSPFLRYNRQIGEFYTRAGLYLYLCOVHDEKAVYLLKD 180
QY 216 LVDGVLALRCLEFSATASSIGPOLRLCOVSGLLALRPSSLRIRTLPMANHLKAAPFL 275
DB 181 LVDGVLALRCLEFSATASSIGPOLRLCOVSGLLALRPSSLRIRTLPMANHLKAAPFL 240
QY 276 TYFGLFQVH 284
DB 241 TYFGLFQVH 249

RESULT 6
AAW29745
ID AAW29745 standard; Protein; 249 AA.
XX
AC AAW29745;
XX
DT 27-OCT-1998 (first entry)
XX

DE	TNF related endothelium proliferative agent protein.	
XX		
KW	TNF: endothelium proliferative agent; TREPA: wound healing; cancer;	
KM	tissue grafting; vascularisation; apoptosis; autoimmune; birth control.	
XX		
OS	Homo sapiens.	
XX		
PM	MO9835061-A2.	
XX		
PD	13-AUG-1998.	
XX		
PE	12-FEB-1998; 98WO-0502859.	
XX		
PR	10-FEB-1998; 98US-0021706.	
XX		
PR	12-FEB-1997; 97US-0798692.	
XX		
PA	(ABBO) ABBOTT LAB.	
XX		
PI	Wiley SR;	
XX		
DR	WPI: 1998-447255/38.	
DR	N-PSDB: AAV47613.	
XX		
PT	Detecting nucleic acid encoding TREPA - useful for diagnosis and	
XX	treatment of autoimmune disease, tumours and inflammation	
XX		
PS	Claim 16; Page 123-4; 142pp: English.	
XX		
CC	The TNF-related endothelium proliferative agent (TREPA), or its	
CC	activators or agonists, are used to treat a deficit of TREPA, e.g. to	
CC	promote wound healing or tissue grafting, by promoting vascularisation,	
CC	also to induce apoptosis for treating cancer and eliminating autoreactive	
CC	T cells, as an adjunct to cancer chemotherapy or antiviral treatment.	
CC	TREPA peptides can also be used to target cytotoxic agents or for	
CC	affinity isolation of the corresponding receptor, the nucleic acid for	
CC	which can be used to transform tumour cells to render them more	
CC	responsive to TREPA and to screen for TREPA mimics.	
CC	Ribozymes, antisense RNA, antibodies or peptides, are used to treat	
CC	TREPA-associated diseases, e.g. tumours and metastases (by inhibiting	
CC	vascularisation), inflammation or a wide range of autoimmune conditions,	
CC	conditions involving abnormal stimulation of epithelial cells (e.g.	
CC	atherosclerosis) for birth control (inhibiting ovulation and placental	
CC	formation) or other angiogenic conditions (e.g. ulcers).	
XX		
CC		
SO	Sequence 249 AA;	
Query Match	84.9%; Score 241; DB 19; Length 249;	
Best Local Similarity	100.0%; Pred. No. 6.1e-213;	
Matches 241: Conservative	0; Mismatches 0; Indels 0; Gaps 0	
OY	44 RRGRRGPGTALVPLALGLALACTGLLAVVSLGSRASLSAOPAEELVAEEDDP	103
Db	9 RRGRRGPGTALVPLALGLALACTGLLAVVSLGSRASLSAOPAEELVAEEDDP	68
OY	104 SELNPOTEESODPADPLNLVLRPRRSAPRGKTRARRATAAHYEVNRPDQDGAQGVND	163
Db	69 SELNPOTEESODPADPLNLVLRPRRSAPRGKTRARRATAAHYEVNRPDQDGAQGVND	128
OY	164 TVSGMEEARINSSPLRKNRQIGEFVTPRAGLYLVCQVNPDEGKAVYLKLDLVDGVLA	223
Db	129 TVSGMEEARINSSPLRKNRQIGEFVTPRAGLYLVCQVNPDEGKAVYLKLDLVDGVLA	188
OY	224 LRCLFEFSSTAASSLGPOLRLCOVSGILALRPGSSLRITRLPMANKAPFLTYFGLFQV	283
Db	189 LRCLFEFSSTAASSLGPOLRLCOVSGILALRPGSSLRITRLPMANKAPFLTYFGLFQV	248
OY	284 H 284	
Db	249 H 249	

ID	AAE00891 standard; Protein: 249 AA.
XX	
AC	AAEC0891;
XX	
DT	04-JUL-2001 (first entry)
XX	
DE	Human TREPA (TNF related endothelium proliferative agent).
XX	
KW	Human; tumour necrosis factor; TNF; angiogenesis; wound healing; TREPA; TNF related endothelium proliferative agent; tumour; metastasis; grafting; vulnerability.
KM	
XX	
OS	Homo sapiens.
FH	
FT	Key Location/Qualifiers Domain 98..249 /label= Extracellular_domain
XX	
PN	US6207642-B1.
PD	
27-MAR-2001.	
PE	26-JUN-1998; 98US-0105343.
XX	
PR	12-FEB-1997; 97US-0798692. 10-FEB-1998; 98US-0021706.
PA	(ABBO) ABBOTT LAB.
XX	
PI	Wiley SR;
XX	
DR	WPI: 2001-280760/29. N-Psdb: AAD04350.
PT	
XX	
PT	Inducing angiogenesis in mammal at desired sites for promoting wound healing, by administering soluble fragment of extracellular domain of tumour necrosis factor related endothelium proliferative agent protein
PT	-
Claim 1; Column 75-76; 53pp; English.	
CC	The present invention relates to extracellular signal molecules, particularly members of tumour necrosis factor (TNF) family molecules designated as TREPA (TNF related endothelium proliferative agent). Soluble biologically active TREPA are used to treat TREPA-associated diseases, tumours or metastases. TREPA is used for inducing angiogenesis in human for promoting wound healing and for vascularising grafted tissue for successful grafting and to promote tissue grafts.
CC	The present amino acid sequence is clone ID #690050 human TREPA.
SQ	Sequence 249 AA:
Query Match	84.9%; Score 241; DB 22; Length 249;
Best Local Similarity	100.0%; Pred. No. 6.1e-213;
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY 44 RRGRRGEPTALLVPALGLATLACTLGLLAVSLGSRASLSAQPQEEIVAEEDDP 103	
DB 9 RRGRGEGETALLVPLALGLTLAGLTLCTGLLAVALSLSRASLSAQEPQDELVAEEDDP 68	
QY 104 SELNPOTESODPADFLNLRLVPRRSAPKGRKTRARRAIAAHVEYNRPDGDGAAGVDG 163	
DB 69 SELNPOTESODPAPFLNLRLVPRRSAPKGRKTRARRAIAAHVEYNRPDGDGAAGVDG 128	
QY 164 TVSGMEEARINSSPLRKNRROIGEIVIRAGIYLIXCOVNHEDEGAAVYLKIDLVDGYLA 223	
DB 129 TVSGMEEARINSSPLRKNRROIGEIVIRAGIYLIXCOVNHEDEGAAVYLKIDLVDGYLA 188	
QY 224 LRCLLEFSATASSLGPOLRCOVSGLLALRPGSSLRIRTLPMAHLKAAPFLTFGFGLFOV 283	
DB 189 LRCLLEFSATASSLGPOLRCOVSGLLALRPGSSLRIRTLPMANLKAAPFLTFGFGLFOV 248	
QY 284 H 284	

Db	249	H 249
RESULT	8	
ID	AAU03499	
AC	AAU03499	standard; Protein: 273 AA.
XX		
XX	AAU03499;	
DT	26-SEP-2001	(first entry)
DE	TWEAK	extracellular domain-containing fusion protein.
XX		
KW	TWEAK	extracellular domain; tumour necrosis factor; TNF; angiogenesis;
KW	ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;	
KW	retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;	
KW	retinosis; uveitis; macular degeneration; arthritis; rheumatism;	
KW	corneal graft neovascularisation; psoriasis; metastatic condition;	
KW	malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;	
KW	preneoplastic condition; myocardial angiogenesis; wound granulation;	
KW	scleroderma; vascular adhesion; telangiectasia; ischaemia; human;	
KW	atherosclerotic plaque neovascularisation; coronary atherosclerosis;	
KW	peripheral atherosclerosis; pDC409-LZ-TWEAK; TWEAK receptor; TWEAK;	
KW	fusion protein.	
XX		
XX	Hom sapiens.	
OS	Synthetic.	
XX		
PN	WO200145730-A2.	
XX		
PD	28-JUN-2001.	
XX		
PE	19-DEC-2000; 2000MO-US34755.	
PR	20-DEC-1999; 99US-0172878.	
PR	10-MAY-2000; 2000US-0203347.	
XX		
PA	(IMMV) IMMUNEX CORP.	
PI	Wiley SR;	
XX	WPI: 2001-417975/44.	
DR	N-PSDB: AAS03964.	
XX		
PT	Modulating angiogenesis in a mammal for treating diseases mediated by	
PT	angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or	
PT	peripheral tissue, by administering antagonist or agonist of TWEAK	
PT	receptor	
XX		
PS	Example 1; Page 41; 46pp; English.	
XX		
CC	The sequence represents a fusion protein encoded by the	
CC	expression vector pDC409-LZ-TWEAK. The fusion protein comprises a	
CC	growth hormone leader, a leucine zipper multimerisation domain, and	
CC	the extracellular domain of human TWEAK. The fusion protein was	
CC	used in the isolation of human TWEAK receptor (TWEAKR)-expressing	
CC	clones from a COS cell human cDNA library. The TWEAK protein is	
CC	a member of the tumour necrosis factor (TNF) family and induces	
CC	angiogenesis. TWEAKR may therefore be used to screen for and	
CC	develop TWEAKR agonists and antagonists for the modulation of	
CC	angiogenesis, to be used in the treatment and diagnosis of human disease	
CC	The disorders mediated by angiogenesis include ocular disorders	
CC	characterised by ocular neovascularisation such as diabetic retinopathy,	
CC	neovascular glaucoma, retinoblastoma, retinopathy of prematurity,	
CC	retrolental fibroplasia, rubeosis, uveitis, macular degeneration and	
CC	corneal graft neovascularisation, and inflammatory diseases such as	
CC	arthritis, rheumatism and psoriasis. Other treatable diseases include	
CC	malignant and metastatic conditions such as sarcomas and carcinomas,	
CC	benign tumours and preneoplastic conditions, myocardial angiogenesis,	
CC	haemophilic joints, scleroderma, vascular adhesions, atherosclerotic	
CC	plaque neovascularisation, telangiectasis, wound granulation, coronary	
CC	atherosclerosis, peripheral atherosclerosis and ischaemia.	

XX	Sequence	273 AA:	72.9%;	Score 207;	DB 22;	Length 273;
XX	Query Match		100.0%;	Pred. No. 1,1e-181;		
XX	Best Local Similarity		0;	Mismatches 0;	Indels 0;	Gaps 0
XX	Matches 207; Conservative					
OY	78 SLGSRASISAOEPQAEELVAEEDODPSELNPOQEEQODPAPFLNRLVRRRSAPGKRTR	137				
Db	67 SLGSRASISAOEPQAEELVAEEDODPSELNPOQEEQODPAPFLNRLVRRRSAPGKRTR	126				
OY	138 ARRAIAAHVEYHPPRGDGAQAGVDGTSGWEFARINSSPLRYNRQIGETLVTRAGLY	197				
Db	127 ARRAIAAHVEYHPPRGDGAQAGVDGTSGWEFARINSSPLRYNRQIGETLVTRAGLY	186				
OY	198 LYCOVHDEGKAVYLKLDLLVDGVLALRCLEEFSSATAASSLGPOLRLCOVSGLLALRPGS	257				
Db	187 LYCOVHDEGKAVYLKLDLLVDGVLALRCLEEFSSATAASSLGPOLRLCOVSGLLALRPGS	246				
OY	258 SLRIRTLPMWHLKKAAPLTYFGLEQVH 284					
Db	247 SLRIRTLPMWHLKKAAPLTYFGLEQVH 273					
XX	RESULT 9					
XX	AAE00895					
XX	AAE00895 standard; Protein; 146 AA.					
XX	AAE00895:					
XX	04-JUL-2001 (first entry)					
XX	Human TREPA (TNF related endothelium proliferative agent) fragment.					
XX	Human: tumour necrosis factor; TNF; angiogenesis; wound healing;					
XX	TREPA: TNF related endothelium proliferative agent; tumour; metastasis;					
XX	grafting; vlnetary.					
XX	Homo sapiens.					
XX	US6207642-B1.					
XX	27-MAR-2001.					
XX	26-JUN-1998; 9805-0105343.					
XX	12-FEB-1997; 9705-0798692.					
XX	10-FEB-1998; 9805-0021706.					
XX	(ABBO) ABBOTT LAB.					
XX	Wiley SR;					
XX	WPI; 2001-280760/29.					
XX	Inducing angiogenesis in mammal at desired sites for promoting wound					
XX	healing, by administering soluble fragment of extracellular domain of					
XX	tumour necrosis factor related endothelium proliferative agent protein					
XX	-					
XX	Example 14; Fig 1; 53pp; English.					
XX	The present invention relates to extracellular signal molecules,					
XX	particularly members of tumour necrosis factor (TNF) family molecules					
XX	designated as TREPA (TNF related endothelium proliferative agent).					
XX	Soluble biologically active TREPA are used to treat TREPA-associated					
XX	diseases, tumours or metastases. TREPA is used for inducing angiogenesis					
XX	in human for promoting wound healing and for vascularising grafted					
XX	tissue for successful grafting and to promote tissue grafts.					
XX	The present amino acid sequence is human TREPA fragment.					
XX	Sequence 146 AA;					

Query Match 51.4%; Score 146; DB 22; Length 146;
 Best Local Similarity 100.0%; Pred. No. 6,4e-126;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 139 RRAIAHYEVPHPGQGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYYL 198
 DB 1 RRAIAHYEVPHPGQGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYYL 60
 QY 199 YCOVHPDEGKAVYLLKDLLVDGVALRLCLEEFSSATAASSLGPQLRLCOVSGLLALRPSS 258
 DB 61 YCOVHPDEGKAVYLLKDLLVDGVALRLCLEEFSSATAASSLGPQLRLCOVSGLLALRPSS 120
 QY 259 LRIRTPMAHLKAAPFLTYFGLFOVH 284
 DB 121 LRIRTPMAHLKAAPFLTYFGLFOVH 146

RESULT 10
 AAM29746
 ID AAM29746 standard; Protein: 189 AA.
 AC AAM29746;
 DT 27-OCT-1998 (first entry)
 DE TNF related endothelium proliferative agent protein 2.
 KM TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
 KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
 OS Homo sapiens.
 PN WO9835061-A2.
 PD 13-AUG-1998.
 PE 12-FEB-1998; 98MO-US02859.
 PR 10-FEB-1998; 98US-0021706.
 PR 12-FEB-1997; 97US-0798692.
 PA (ABBO) ABBOTT LAB.
 PI Wiley SR;
 PS WPI; 1998-447255/38.
 DR Detecting nucleic acid encoding TREPA - useful for diagnosis and
 PT treatment of autoimmune disease, tumours and inflammation
 PS Claim 16; Page 125-6; 142pp; English.

CC The TNF-related endothelium proliferative agent (TREPA), or its
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
 CC promote wound healing or tissue grafting, by promoting vascularisation,
 CC also to induce apoptosis for treating cancer and eliminating autoreactive
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
 CC TREPA peptides can also be used to target cytotoxic agents or for
 CC affinity isolation of the corresponding receptor, the nucleic acid for
 CC which can be used to transform tumour cells to render them more
 CC responsive to TREPA and to screen for TREPA mimics.
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
 CC vascularisation), inflammation or a wide range of autoimmune conditions,
 CC conditions involving abnormal stimulation of epithelial cells (e.g.
 CC atherosclerosis), for birth control (inhibiting ovulation and placental
 CC formation) or other angiogenic conditions (e.g. ulcers).

SQ Sequence 189 AA;

Query Match 50.4%; Score 143; DB 19; Length 189;
 Best Local Similarity 100.0%; Pred. No. 4,6e-123;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 IAAHYEVPHPGQGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLXCQ 201
 DB 47 IAAHYEVPHPGQGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLXCQ 106
 QY 202 VHPDEGKAVYLLKDLLVDGVALRLCLEEFSSATAASSLGPQLRLCOVSGLLALRPSSLRI 261
 DB 107 VHPDEGKAVYLLKDLLVDGVALRLCLEEFSSATAASSLGPQLRLCOVSGLLALRPSSLRI 166
 QY 262 RTLPMAHLKAAPFLTYFGLFOVH 284
 DB 167 RTLPMAHLKAAPFLTYFGLFOVH 189

RESULT 11
 AAE00892
 ID AAE00892 standard; Protein: 189 AA.
 AC AAE00892;
 DT 04-JUL-2001 (first entry)
 DE Human UL4flag TREPA soluble construct.
 KM Human; tumour necrosis factor; TNF; angiogenesis; wound healing; tumour;
 KW TREPA; TNF related endothelium proliferative agent; metastasis; grafting;
 KW vulnereary; HUVEC; human umbilical vein endothelial cell; UL4flag.
 OS Homo sapiens.
 PN US6207642-B1.
 PD 27-MAR-2001.
 PE 26-JUN-1998; 98US-0105343.
 PR 12-FEB-1997; 97US-0798692.
 PR 10-FEB-1998; 98US-0021706.
 PA (ABBO) ABBOTT LAB.
 PI Wiley SR;
 PS WPI; 2001-280760/29.

PT Inducing angiogenesis in mammal at desired sites for promoting wound
 PT healing, by administering soluble fragment of extracellular domain of
 PT tumor necrosis factor related endothelium proliferative agent protein
 PS -

CC Example 2; Column 75-78; 53pp; English.

CC The present invention relates to extracellular signal molecules,
 CC particularly members of tumour necrosis factor (TNF) family molecules
 CC designated as TREPA (TNF related endothelium proliferative agent).
 CC Soluble biologically active TREPA are used to treat TREPA-associated
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
 CC in human for promoting wound healing and for vascularising grafted
 CC tissue for successful grafting and to promote tissue grafts. The present
 CC amino acid sequence is human UL4flag TREPA soluble construct. This
 CC sequence which is a biologically active molecule is capable of inducing
 CC proliferation in HUVEC (human umbilical vein endothelial cells) cells.

SQ Sequence 189 AA;

Query Match 50.4%; Score 143; DB 22; Length 189;
 Best Local Similarity 100.0%; Pred. No. 4,6e-123;
 Matches 143; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 142 IAAHYEVPHPGQGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLXCQ 201
 DB 47 IAAHYEVPHPGQGAQAGVDGTVSGWEARINSSPLRYNRQIGEFIVTRAGLYLXCQ 106

```

Oy 202 VHFDEGKAVYKLDLVDGYALRCLEEFSSATAASSLGQRLRCQVSGLLALPGSSLR 261
    |||
Db 107 VHFDEGKAVYKLDLVDGYALRCLEEFSSATAASSLGQRLRCQVSGLLALPGSSLR 166
Oy 262 RFLPWAHLKAAPFLTYFGLFOVH 284
    |||
Db 167 RFLPWAHLKAAPFLTYFGLFOVH 189

RESULT 12
AAW93590
ID AAW93590 standard: Protein; 208 AA.
AC AAW93590;
XX
XX 18-JUN-1999 (first entry)
DT
XX
XX Human TNRL3 protein.
DE
XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW developmental abnormality; gestational abnormality; prostate cancer;
KW APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW apoptosis; human.
XX
XX Homo sapiens.
OS
XX
XX W09911791-A2.
PN
XX
XX 11-MAR-1999.
PD
XX
XX 04-SEP-1998; 98WO-US18393.
PF
XX
XX 05-SEP-1997; 97US-0924634.
PR
XX
XX (UNITM ) UNITV WASHINGTON.
PA
XX
XX Chaudhary PM;
PI
XX
XX WPI; 1999-205191/17.
DR
XX
XX N-PSDB; AAX23424.
PT
XX
XX New Tumour Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
PS
XX
XX Claim 40; Fig 13A; 156pp; English.
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the changer in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
CC
XX Sequence 208 AA:
Query Match 37.7%; Score 107; DB 20; Length 208;

```

```

Best Local Similarity 99.5%; Pred. No. 5.8e-90;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Oy 77 VSLGSRASLSAEPPOEELVAEEDDPSELNPQTESODPAPEFLNRLVPRRSAPKGRKT 136
    |||
Db 1 VSLGSRASLSAEPPOEELVAEEDDPSELNPQTESODPAPEFLNRLVPRRSAPKGRKT 60
Oy 137 RARRAIAHYEVHPRGODGAGVGTSGWEARINSSPLRYNRQIGEIFVRAGLY 196
    |||
Db 61 RARRAIAHYEVHPRGODGAGVGTSGWEARINSSPLRYNRQIGEIFVRAGLY 120
Oy 197 YLYCOVHFDEGKAVYKLDLVDGYALRCLEEFSSATAASSLGQRLRCQVSGLLALRG 256
    |||
Db 121 YLYCOVHFDEGKAVYKLDLVDGYALRCLEEFSSATAASSLGQRLRCQVSGLLALRG 180
Oy 257 SSLRITRFLPWAHLKAAPFLTYFGLFOVH 284
    |||
Db 181 SSLRITRFLPWAHLKAAPFLTYFGLFOVH 208

RESULT 13
AAW93591
ID AAW93591 standard: Protein; 211 AA.
AC AAW93591;
XX
XX 18-JUN-1999 (first entry)
DT
XX
XX Mouse TNRL3 protein.
DE
XX
XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KW developmental abnormality; gestational abnormality; prostate cancer;
KW APO6; APO8; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KW apoptosis; mouse.
XX
XX Mus sp.
OS
XX
XX W09911791-A2.
PN
XX
XX 11-MAR-1999.
PD
XX
XX 04-SEP-1998; 98WO-US18393.
PF
XX
XX 05-SEP-1997; 97US-0924634.
PR
XX
XX (UNITM ) UNITV WASHINGTON.
PA
XX
XX Chaudhary PM;
PI
XX
XX WPI; 1999-205191/17.
DR
XX
XX N-PSDB; AAX23425.
PT
XX
XX New Tumour Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
PS
XX
XX Claim 40; Fig 13B; 156pp; English.
CC This invention describes isolated Tumor Necrosis Factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of
CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the changer in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active

```

CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO4 was transfected to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.

XX
SQ Sequence 211 AA;

Query Match 16.2%; Score 46; DB 20; Length 211;
Best Local Similarity 100.0%; Pred. No. 6.1e-34;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 239 GPOLRLGVSGSLALRPSSLRITLPMAHLKAPFLTYGLPQVH 284
Db 166 GPOLRLGVSGSLALRPSSLRITLPMAHLKAPFLTYGLPQVH 211

RESULT 14

AAW47524
ID AAW47524 standard; Protein: 225 AA.

AC AAW47524;

DT 21-JUL-1998 (first entry)

DE Mus musculus tumour necrosis factor related ligand (TRELL).

XX
XX TRELL; tumour necrosis factor related ligand; tnf; treatment;
KW cancer; autoimmune disease; immune system; stimulation; suppression;
KW graft rejection.

XX
OS Mus musculus.

XX
FH Key Location/Qualifiers

FT Domain 1..21 /note="hydrophobic, transmembrane domain"

XX
PN WO9805783-A1.

XX
PD 12-FEB-1998.

XX
PF 07-AUG-1997; 97WO-US13945.

XX
PR 18-MAR-1997; 97US-0040820.

XX
PR 07-AUG-1996; 96US-0023541.

XX
PR 18-OCT-1996; 96US-0028515.

XX
PA (BIOJ) BIOGEN INC.

XX
PI (UYGE-) UNIV GENEVA FACULTY MEDICINE.

XX
PI Browning JL, Chicheportiche Y;

XX
DR WPI; 1998-145619/13.

XX
DR N-PDB; AAV18599.

XX
XX Claim 12; Pages 48-50; 69pp; English.

CC The sequence is that of mouse tumour necrosis factor related
CC ligand (TRELL). TRELL or active fragments can be included with a
CC carrier in pharmaceutical compositions to treat cancer, autoimmune
CC diseases or immune responses to tissue grafts, or to stimulate or
CC suppress the immune system. It is useful to screen for TRELL
CC receptors, by labelling with a detectable label and screening
CC compositions for binding. Agents interfering with TRELL-receptor
CC binding can also be screened for, can then be administered,
CC optionally with interferon- gamma, to induce cell death or
CC treat, suppress or alter immune responses (especially involving human
CC adenocarcinoma cells) involving a signal pathway between TRELL and its

CC receptor. It's coding sequence can be used in gene therapy for
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
CC autoimmune and inflammatory diseases or inherited genetic disorders,
CC by introducing into cells, and expressing, therapeutically effective
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
CC It may also be of use in the preparation of probe primers for
CC screening natural/synthetic DNAs for TRELL-encoding sequences
CC and for antisense therapy.

XX
SQ Sequence 225 AA;

Query Match 11.3%; Score 32; DB 19; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.7e-21;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 139 RRAIAHYEVHPRGCGAAGVDTYSGWEE 170
Db 80 RRAIAHYEVHPRGCGAAGVDTYSGWEE 111

RESULT 15

AAB07527
ID AAB07527 standard; Protein: 225 AA.

AC AAB07527;

DT 20-OCT-2000 (first entry)

DE Amino acid sequence of a soluble recombinant murine TWEAK protein.

XX
XX TWEAK protein; immunological disorder; immune response; inflammation;
KW TWEAK blocking agent; autoimmune disease; organ transplant rejection;

XX
KW Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.

XX
OS Mus sp.

XX
PN WO200042073-A1.

XX
PD 20-JUL-2000.

XX
PE 14-JAN-2000; 2000WO-US01044.

XX
PR 15-JAN-1999; 99US-0116168.

XX
PA (BIOJ) BIOGEN INC.

XX
PI Renner P;

XX
DR WPI; 2000-476036/41.

XX
XX Preventing and treating immune responses using modulators, especially
PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for

XX
PT treating e.g. inflammation and graft versus host disease -

XX
PS Disclosure; Fig 1; 45pp; English.

CC The present sequence represents a TWEAK protein. The specification
CC describes a method for preventing or treating an immunological
CC disorder and/or inhibiting an immune response in an animal. The
CC method comprises administering a TWEAK blocking agent. The method may
CC be used for preventing and treating immune disorders associated with
CC inappropriate expression and/or activity of TWEAK. These disorders
CC include autoimmune diseases, acute and chronic inflammation, organ
CC transplant rejection, graft-versus-host disease (GVHD), lymphoid cell
CC malignancies, septic and other forms of shock, loss of immune
CC responsiveness (as seen in human immunodeficiency virus (HIV)
CC infections) and failure of the immune response to tumour growth.

XX
SQ Sequence 225 AA;

Query Match 11.3%; Score 32; DB 21; Length 225;
Best Local Similarity 100.0%; Pred. No. 4.7e-21;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 RRAIAHYEVHPRPGDGAQAGVDGTVSGMEE 170
 |||||
DB 80 RRAIAHYEVHPRPGDGAQAGVDGTVSGMEE 111

RESULT 16
AAG01265
ID AAG01265 standard; Protein; 58 AA.
XX
AC AAG01265;
XX
DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 5346.
XX
DE Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EPI033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PS (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-500381/45.
DR N-PSDB: AAC01271.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX
PS Claim 13; SEQ ID 5346; 71pp + CD-ROM; English.
XX
CC The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
SQ Sequence 58 AA:

Query Match 3.2%; Score 9; DB 21; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSLDPEIS 9
 |||||
DB 46 MSLDPEIS 54

RESULT 17
AAG01266
ID AAG01266 standard; Protein; 58 AA.
XX
AC AAG01266;
XX

DT 06-OCT-2000 (first entry)
XX
DE Human secreted protein, SEQ ID NO: 5347.
XX
KM Human, 5' EST; expressed sequence tag; secreted protein; cDNA isolation;
KM gene therapy; chromosome mapping.
XX
OS Homo sapiens.
XX
PN EPI033401-A2.
XX
PD 06-SEP-2000.
XX
PF 21-FEB-2000; 2000EP-0200610.
XX
PR 26-FEB-1999; 99US-0122487.
XX
PS (GEST) GENSET.
XX
PI Dumas Milne Edwards J, Duclert A, Giordano J;
XX
DR WPI: 2000-500381/45.
DR N-PSDB: AAC01272.
XX
PT New nucleic acid that is a 5' expressed sequence tag (5' EST) for
PT obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for
PT diagnostic, forensic, gene therapy and chromosome mapping procedures -
XX
XX
PS Claim 13; SEQ ID 5347; 71pp + CD-ROM; English.
XX
CC The present sequence is a polypeptide encoded by one of a large number
CC of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs
CC were prepared from total human RNAs or polyA+ RNAs derived from 30
CC different tissues. EST sequences usually correspond mainly to the 3'
CC untranslated region (UTR) of the mRNA because they are often obtained
CC from oligo-dT primed cDNA libraries. Such ESTs are not well suited for
CC isolating cDNA sequences derived from the 5' ends of mRNAs and even in
CC those cases where longer cDNA sequences have been obtained, the full 5'
CC UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'
CC ends and can therefore be used to obtain full length cDNAs and genomic
CC DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and
CC chromosome mapping procedures. They are used to obtain upstream
CC regulatory sequences and to design expression and secretion vectors.
XX
SQ Sequence 58 AA:

Query Match 3.2%; Score 9; DB 21; Length 58;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MSLDPEIS 9
 |||||
DB 46 MSLDPEIS 54

RESULT 18
AAY14906
ID AAY14906 standard; protein; 748 AA.
XX
AC AAY14906;
XX
DT 25-OCT-1999 (first entry)
XX
DE Extended amino acid sequence for GVs-9.
XX
KM Mycobacterium vaccae protein; antigen; T cell activation; cytokine;
KM dendritic cell maturation; infectious disease; immune disorder; cancer;
KM respiratory system; mycobacterial infection; allergy; tuberculosis;
KM leprosy; sarcoidosis; lung cancer; asthma; skin disorder; psoriasis;
KM dermatitis; eczema; alopecia areata; skin cancer; basal carcinoma;
KM squamous cell carcinoma; melanoma.
XX
OS Mycobacterium vaccae.

XX MO9932634-A2.
 PN 01-JUL-1999.
 XX
 PD 23-DEC-1998; 98WO-N200189.
 XX
 PF 04-DEC-1998; 98US-0205426.
 PR 23-DEC-1997; 97US-0996624.
 PR 23-DEC-1997; 97US-0997080.
 PR 23-DEC-1997; 97US-0997362.
 PR 11-JUN-1998; 98US-0095855.
 PR 17-SEP-1998; 98US-0156181.
 XX
 PA (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PI Prestidge RL, Skinner MA, Tan P, Visser ES, Watson J;
 XX
 DR WPI; 1999-430163/36.
 DR N-PSDB; AA211368.
 XX
 XX Enhancing immune response to an antigen
 PT
 PS Claim 1; Page 209-210; 243pp; English.
 XX
 PS The invention provides heat-killed Mycobacterium vaccae, or recombinant
 CC M. vaccae proteins. The M. vaccae proteins may be employed to activate
 CC T cells and natural killer cells, to stimulate the production of
 CC cytokines, to enhance the expression of co-stimulatory molecules on
 CC dendritic cells and monocytes, and to enhance dendritic cell maturation
 CC and function. The proteins can be expressed by standard recombinant
 CC methodology. Pharmaceutical compositions comprising the proteins or
 CC nucleic acid sequences encoding the proteins can be used for the
 CC treatment, prevention, and detection of disorders including infectious
 CC diseases, immune disorders and cancer. In particular, the compounds and
 CC methods are used for treatment of diseases of the respiratory system,
 CC such as mycobacterial infections, asthma, allergies, tuberculosis,
 CC leprosy, sarcoidosis and lung cancers, and disorders of the skin such as
 CC psoriasis, atopic dermatitis, eczema, allergic contact dermatitis,
 CC alopecia areata, and skin cancers such as basal carcinoma, squamous cell
 CC carcinoma and melanoma.
 CC
 XX
 SQ Sequence 748 AA;
 Query Match 3.2%; Score 9; DB 20; Length 748;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 ALGIGLALA 68
 |||||||||
 Db 282 ALGIGLALA 290
 RESULT 19
 ABB73512
 ID ABB73512 standard; Protein; 749 AA.
 XX
 AC ABB73512;
 XX
 DT 08-APR-2002 (first entry)
 XX
 DE M vaccae GVs-9 protein SEQ ID NO: 154.
 XX
 DE Skin disorder; psoriasis; atopic dermatitis; allergic contact dermatitis;
 KW alopecia areata; skin cancer; Mycobacterium vaccae; melanoma; cytostatic;
 KW antipsoriatic; dermatological; antiinflammatory; antiallergic;
 KW Th2 immune response; immunomodulatory.
 XX
 OS Mycobacterium vaccae.
 XX
 PI US6328978-B1.
 XX
 PD 11-DEC-2001.

XX 02-JUN-1999; 99US-0324542.
 XX
 PF 23-DEC-1997; 97US-0997080.
 XX
 PR (GENE-) GENESIS RES & DEV CORP LTD.
 XX
 PA Watson JD, Tan PLJ, Prestidge R;
 XX
 PI WPI; 2002-138361/18.
 DR N-PSDB; ABL36274.
 XX
 XX Inhibiting skin inflammation associated with skin disorder e.g.
 PT psoriasis, by administering composition comprising delipidated and
 PT deglycolipidated Mycobacterium vaccae cells or Mycobacterium vaccae
 PT culture filtrate -
 XX
 PS Example 6; Column 161-164; 116pp; English.
 XX
 CC The present invention relates to a method of inhibiting skin inflammation
 CC associated with a skin disorder selected from psoriasis, atopic
 CC dermatitis and allergic contact dermatitis, which involves administering
 CC a composition containing delipidated and deglycolipidated Mycobacterium
 CC vaccae cells or M. vaccae culture filtrate. The skin disorder to be
 CC treated may also include alopecia areata, and skin cancers such as basal
 CC cell carcinoma, squamous cell carcinoma and melanoma. The composition
 CC acts by inhibiting the Th2 immune response. The present sequence is a
 CC protein described in the exemplification of the invention.
 CC
 XX
 SQ Sequence 749 AA;
 Query Match 3.2%; Score 9; DB 23; Length 749;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 60 ALGIGLALA 68
 |||||||||
 Db 282 ALGIGLALA 290
 RESULT 20
 AAU51863
 ID AAU51863 standard; Protein; 54 AA.
 XX
 AC AAU51863;
 XX
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #12759.
 XX
 DE SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 20-APR-2001; 2001WO-US12865.
 XX
 PR 21-APR-2000; 2000US-199047P.
 PR 02-JUN-2000; 2000US-208841P.
 PR 07-JUL-2000; 2000US-216747P.
 XX
 PA (CORI-) CORIXA CORP.
 XX
 PI Skelky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.

DR N-PSDB; AAS59552.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
PS Example 1; SEQ ID NO 13058; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 54 AA:

Query Match 2.8%; Score 8; DB 22; Length 54;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 11 RLPLPRS 18
DB 21 RLPLPRS 28

RESULT 21
ABBA42637
ID ABBA42637 standard; Peptide: 55 AA.
XX
AC ABBA42637;
XX
DT 04-FEB-2002 (first entry)
XX
DE Peptide #10143 encoded by human foetal liver single exon probe.
XX
KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
OS Homo sapiens.
XX
PN WO200157277-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00669.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483447/52.

XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human fetal liver -
XX
PS Claim 27; SEQ ID NO 35272; 639pp + sequence listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for
CC measuring human gene expression in a sample derived from human foetal
CC liver. The single exon nucleic acid probes may be used for predicting,
CC measuring and displaying gene expression in samples derived from human
CC fetal liver. The present sequence is a peptide encoded by a single exon
CC nucleic acid probe of the invention.
CC Note: The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 55 AA:

Query Match 2.8%; Score 8; DB 22; Length 55;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 61 LGLGLALA 68
DB 12 LGLGLALA 19

RESULT 22
ABB25983
ID ABB25983 standard; Protein: 55 AA.
XX
AC ABB25983;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #7982 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US00666.
XX
PR 04-FEB-2000; 2000US-0180312.
XX
PR 26-MAY-2000; 2000US-0207456.
XX
PR 30-JUN-2000; 2000US-0608408.
XX
PR 03-AUG-2000; 2000US-0632366.
XX
PR 21-SEP-2000; 2000US-0234687.
XX
PR 27-SEP-2000; 2000US-0236359.
XX
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15; SEQ ID NO 27753; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABBA1305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.

CC By measuring gene expression, the probes are useful for predicting,
 CC diagnosing, grading, staging, monitoring and prognosing diseases of the
 CC human heart and vascular system e.g. cardiovascular disease,
 CC hypertension, cardiac arrhythmias and congenital heart disease.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.

XX
 CC Sequence 55 AA;

Query Match 2.8%; Score 8; DB 22; Length 55;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68
 Db 12 LGIGLALA 19

RESULT 23
 AAM63528

ID AAM63528 standard; Protein; 55 AA.

AC AAM63528;

XX 05-NOV-2001 (first entry)

DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35633.

XX Human; brain expressed exon; gene expression analysis; probe;

KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;

XX epilepsy; cancer.

OS Homo sapiens.

PN W0200157275-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001MO-US00667.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-483446/52.

DR Single exon nucleic acid probes for analyzing gene expression in human

XX brains -

PT Example 4; SEQ ID NO: 35633; 650pp + Sequence Listing; English.

PS The present invention provides a number of single exon nucleic acid

XX probes which are derived from genomic sequences expressed in the human

CC brain. They can be used to measure gene expression in brain cell samples,

CC which may enable the diagnosis and improved treatment of nervous system

CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,

CC epilepsy and cancers. The present sequence is a protein encoded by one of

CC the probes of the invention.

XX Sequence 55 AA;

XX Query Match 2.8%; Score 8; DB 22; Length 55;

XX Best Local Similarity 100.0%; Pred. No. 13;

XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68
 Db 12 LGIGLALA 19

RESULT 24
 AAM76342

ID AAM76342 standard; Protein; 55 AA.

XX AAM76342;

XX 06-NOV-2001 (first entry)

DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36648.

XX Human; bone marrow expressed exon; gene expression analysis; probe;

KM microarray; cancer; leukaemia; lymphoma; myeloma.

XX Homo sapiens.

PN W0200157276-A2.

XX 09-AUG-2001.

PD 30-JAN-2001; 2001MO-US00668.

XX 04-FEB-2000; 2000US-0180312.

PR 26-MAY-2000; 2000US-0207456.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-0234687.

PR 27-SEP-2000; 2000US-0236359.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488900/53.

DR Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human bone marrow -

PT Example 4; SEQ ID NO: 36648; 658pp + Sequence Listing; English.

XX The present invention provides a number of single exon nucleic acid

CC probes which are derived from genomic sequences expressed in the human

CC bone marrow. They can be used to measure gene expression in bone marrow

CC samples, which may enable the improved diagnosis and treatment of cancers

CC such as lymphoma, leukaemia and myeloma. The present sequence is a

CC protein encoded by one of the probes of the invention.

XX Sequence 55 AA;

XX Query Match 2.8%; Score 8; DB 22; Length 55;

XX Best Local Similarity 100.0%; Pred. No. 13;

XX Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68

Db 12 LGIGLALA 19

RESULT 25
 AAM20858

ID AAM20858 standard; Protein; 55 AA.

XX AAM20858;

XX 12-OCT-2001 (first entry)

XX Peptide #7292 encoded by probe for measuring cervical gene expression.

[illegible]

PF	30-JAN-2001;	2001WO-US00663.
XX		
PR	04-FEB-2000;	2000US-0180312.
PR	26-MAY-2000;	2000US-0207456.
PR	30-JUN-2000;	2000US-0608408.
PR	03-AUG-2000;	2000US-0632366.
PR	21-SEP-2000;	2000US-0234687.
PR	27-SEP-2000;	2000US-0236359.
PR	04-OCT-2000;	2000GB-0024263.
XX		
PA	(MOLE-)	MOLECULAR DYNAMICS INC.
XX		
PI	Penn SG,	Hanzel DK, Chen W, Rank DR;
XX		
DR	WPI:	2001-488897/53.
XX		
PT	Human genome-derived single exon nucleic acid probes useful for	
XX	analyzing gene expression in human placenta -	
PS	Claim 27; SEQ ID NO 36720;	654pp; English.
XX		
CC	The present invention relates to single exon nucleic acid probes (SENP;	
CC	see AAI3115-AAI57546). The present sequence is a peptide encoded by one	
CC	such probe. The probes are useful for producing a microarray for	
CC	predicting, measuring and displaying gene expression in samples derived	
CC	from human placenta. The probes are useful for antenatal diagnosis of	
CC	human genetic disorders.	
XX		
SQ	Sequence	55 AA;
	Query Match	2.8%; Score 8; DB 22; Length 55;
	Best Local Similarity	100.0%; Pred. No. 13;
	Matches	8; Conservative 0; Mismatches 0; Indels 0; Gaps 0
Oy	61 LGLGLALA 68	
Db	12 LGLGLALA 19	
RESULT 27		
ABG45635		
ID	ABG45635 standard; Peptide; 55 AA.	
XX		
AC	ABG45635;	
XX		
DT	19-AUG-2002 (first entry)	
XX		
DE	Human peptide encoded by genome-derived single exon probe SEQ ID 35300.	
XX		
KM	Human; single exon probe; asthma; lung cancer; COPD; ILD;	
KM	chronic obstructive pulmonary disease; interstitial lung disease;	
KM	familial idiopathic pulmonary fibrosis; neurofibromatosis;	
KM	tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;	
KM	Hernanky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;	
KM	pulmonary histiocytosis; lymphangioleiomyomatosis; Kartagener syndrome;	
KM	pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;	
KM	primary ciliary dyskinesia; pulmonary hypertension;	
XX	hyaline membrane disease.	
OS		
XX	Homo sapiens.	
XX		
PN	WO200186003-A2.	
XX		
PD	15-NOV-2001.	
XX		
PF	30-JAN-2001;	2001WO-US00665.
XX		
PR	04-FEB-2000;	2000US-180312P.
PR	26-MAY-2000;	2000US-207456P.
PR	30-JUN-2000;	2000US-0608408.
PR	03-AUG-2000;	2000US-0632366.
PR	21-SEP-2000;	2000US-234687P.
PR	27-SEP-2000;	2000US-236359P.


```
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488899/53.
XX
DR
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
XX
PS Claim 15; SEQ ID NO 28660; 530pp; English.
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease.
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 65 AA;
XX
Query Match 2.8%; Score 8; DB 22; Length 65;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 LGLGLALA 68
|||
DB 22 LGLGLALA 29
XX
RESULT 30
AAM64984
ID AAM64984 standard; Protein: 65 AA.
XX
AC AAM64984;
XX
XX 05-NOV-2001 (first entry)
XX
DT Human brain expressed single exon probe encoded protein SEQ ID NO: 37089.
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 37089.
XX
XX Human; brain expressed exon; gene expression analysis; probe:
KW microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KW epilepsy; cancer.
XX
XX Homo sapiens.
XX
OS
XX WO200157275-A2.
XX
PN
XX 09-AUG-2001.
XX
PD
XX 30-JAN-2001; 2001WO-US00667.
XX
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
```

```
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX
DR
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
XX
PS Example 4; SEQ ID NO: 37089; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 65 AA;
XX
Query Match 2.8%; Score 8; DB 22; Length 65;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 61 LGLGLALA 68
|||
DB 22 LGLGLALA 29
XX
RESULT 31
AAM77706
ID AAM77706 standard; Protein: 65 AA.
XX
AC AAM77706;
XX
XX 06-NOV-2001 (first entry)
XX
DT Human bone marrow expressed probe encoded protein SEQ ID NO: 38012.
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 38012.
XX
KW Human; bone marrow expressed exon; gene expression analysis; probe:
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
OS
XX WO200157276-A2.
XX
PN
XX 09-AUG-2001.
XX
PD
XX 30-JAN-2001; 2001WO-US00668.
XX
PF
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX
DR
XX
PT Human genome-derived single exon nucleic acid probes useful for
PT analyzing gene expression in human bone marrow -
XX
PS Example 4; SEQ ID NO: 38012; 658pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
```

CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.

SQ Sequence 65 AA;

Query Match 2.8%; Score 8; DB 22; Length 65;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68
| | | | | | | |
Db 22 LGIGLALA 29

RESULT 32

AAM21621
ID AAM21621 standard; Protein; 65 AA.

XX AAM21621;

XX 12-OCT-2001 (first entry)

DE Peptide #8055 encoded by probe for measuring cervical gene expression.

XX Probe: human; microarray; gene expression; cervical epithelial cell;

KM cervical cancer.

XX Homo sapiens.

XX WO200157278-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00670.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488901/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human cervical epithelial cells -

XX Claim 27; SEQ ID No 26447; 487pp; English.

XX The present invention relates to human single exon nucleic acid probes

XX (SENPs: see AAI10068-AA128459). The present sequence is a peptide encoded

XX by one such probe. The SENPs are derived from human HeLa cells. The SENPs

XX can be used to produce a single exon microarray, which can be used for

XX measuring human gene expression in a sample derived from human cervical

XX epithelial cells. By measuring gene expression, the probes are therefore

XX useful in grading and/or staging of diseases of the cervix, notably

XX cervical cancer.

XX Note: The sequence data for this patent did not form part of the printed

XX specification, but was obtained in electronic format directly from WIPO

XX at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 65 AA;

Query Match 2.8%; Score 8; DB 22; Length 65;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68
| | | | | | | |
Db 22 LGIGLALA 29

RESULT 33

AAM37923
ID AAM37923 standard; Protein; 65 AA.

XX AAM37923;

XX 17-OCT-2001 (first entry)

DE Peptide #11960 encoded by probe for measuring placental gene expression.

XX Probe: microarray; human; placenta; antenatal diagnosis;

KM genetic disorder.

XX Homo sapiens.

XX WO200157272-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001WO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0207456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0236359.

XX 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-48897/53.

XX Human genome-derived single exon nucleic acid probes useful for

XX analyzing gene expression in human placenta -

XX Claim 27; SEQ ID No 38192; 654pp; English.

XX The present invention relates to single exon nucleic acid probes (SENPs:

XX see AAI31315-AA157546). The present sequence is a peptide encoded by one

XX such probe. The probes are useful for producing a microarray for

XX predicting, measuring and displaying gene expression in samples derived

XX from human placenta. The probes are useful for antenatal diagnosis of

XX human genetic disorders.

XX Sequence 65 AA;

Query Match 2.8%; Score 8; DB 22; Length 65;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68
| | | | | | | |
Db 22 LGIGLALA 29

RESULT 34

ABG46737
ID ABG46737 standard; Peptide; 65 AA.

XX ABG46737;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 36402.

XX Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW Chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagazer syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX Homo sapiens.
 OS WO200186003-A2.
 FN 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US00665.
 PF 04-FEB-2000; 2000US-180312P.
 PR 26-MAY-2000; 2000US-207456P.
 PR 30-JUN-2000; 2000US-0608408.
 PR 03-AUG-2000; 2000US-0632366.
 PR 21-SEP-2000; 2000US-234687P.
 PR 27-SEP-2000; 2000US-236359P.
 PR 04-OCT-2000; 2000GB-0024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA Penn SC, Hanzel DK, Chen W, Rank DR;
 PI WPI; 2002-114183/15.
 XX Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples -
 PS Claim 27; SEQ ID No 36402; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of
 CC probes; the novel set of probes which hybridise at high stringency to a
 CC nucleic acid expressed in the human lung; measuring gene expression in a
 CC sample derived from human lung, comprising (a) contacting the array with
 CC a collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of
 CC the array; identifying exons in a eukaryotic genome, comprising
 CC (a) algorithmically predicting at least one exon from genomic sequences
 CC of the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene
 CC expression analysis, and for identifying exons in a gene, particularly
 CC using human lung derived mRNA and for the study of lung diseases
 CC such as asthma, lung cancer, chronic obstructive pulmonary disease
 CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary
 CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,
 CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary
 CC haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis,
 CC pulmonary alveolar proteinosis, Karagazer syndrome, fibrocystic
 CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension
 CC and hyaline membrane disease. The present sequence is a peptide/protein

CC encoded by a single exon probe of the invention.
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp://ipo.int/pub/published_pct_sequences.
 CC XX
 SO Sequence 65 AA:
 Query Match 2.8%; Score 8; DB 23; Length 65;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 61 LGLGLALA 68
 |||||
 DB 22 LGLGLALA 29
 RESULT 35
 ID ABG03723 standard; Protein: 69 AA.
 XX ABG03723;
 AC ABG03723;
 XX 13-FEB-2002 (first entry)
 DT
 XX Novel human diagnostic protein #3714.
 DE
 XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 XX Homo sapiens.
 OS WO200175067-A2.
 FN 11-OCT-2001.
 XX 30-MAR-2001; 2001WO-US08631.
 PF 31-MAR-2000; 2000US-0540217.
 PR 23-AUG-2000; 2000US-0649167.
 PR (HYSE-) HYSEQ INC.
 PA Drmanac RT, Liu C, Tang YT;
 PI WPI; 2001-639362/73.
 DR N-PSDB; AAS67910.
 XX New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 PS Claim 20; SEQ ID No 34082; 103pp; English.
 XX The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG0377 represent novel human
 CC diagnostic amino acid sequences of the invention.

CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 84 AA;
Query Match 2.8%; Score 8; DB 22; Length 84;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 259 LRIRLPLW 266
| | | | | | | |
DB 34 LRIRLPLW 41
RESULT 38
ABG20259
ID ABG20259 standard; Protein; 110 AA.
XX
AC ABG20259;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #20250.
XX
KM Human; chromosome mapping; gene mapping; gene therapy; forensic;
KM food supplement; medical imaging; diagnostic; genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200175067-A2.
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
XX
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
XX
DR N-PSDB; AAS84446.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -
XX
PS Claim 20; SEQ ID NO 50618; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human
CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 110 AA;
Query Match 2.8%; Score 8; DB 22; Length 110;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 61 LGLGLALA 68
| | | | | | | |
DB 44 LGLGLALA 51
RESULT 39
AAO08094
ID AAO08094 standard; Protein; 117 AA.
XX
AC AAO08094;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 21986.
XX
DE Human cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US04927.
XX
PR 28-FEB-2000; 2000US-0515126.
XX
PR 18-MAY-2000; 2000US-0577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI: 2001-514838/56.
XX
DR N-PSDB; AAI88025.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing
PT diagnosing and treating e.g. leukaemia, inflammation and immune
PT disorders -
XX
PS Claim 20; SEQ ID NO 21986; 1399bp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 117 AA;
Query Match 2.8%; Score 8; DB 22; Length 117;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 61 LGLGLALA 68
| | | | | | | |

Db 54 LGIGLALA 61

RESULT 40
ABB67712
ID ABB67712 standard; Protein: 184 AA.

XX ABB67712;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 29928.

XX Drosophila: developmental biology; cell signalling; insecticide;
KM pharmaceutical.

XX Drosophila melanogaster.

XX MO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001MO-US09231.

XX 23-MAR-2000; 2000US-191637P.

PR 11-JUL-2000; 2000US-0614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI: 2001-656860/75.

DR N-PSDB; ABL11815.

XX New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -

XX Disclosure: SEQ ID NO 29928; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 184 AA;

Query Match 2.8%; Score 8; DB 22; Length 184;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 72 LLAAVSL 79

DB 16 LLAAVSL 23

Search completed: March 31, 2003, 14:13:36
Job time : 71 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 14:07:46 ; Search time 39 Seconds
(Without alignments)
700.056 Million cell updates/sec

Title: US-09-245-198A-4
Perfect score: 284
Sequence: 1 MSLLDFEISARRLPLPSRLG.....PMAHLKAAPFLTYGLFQVH 284

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size: 0

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 65 summaries

Database : PIR_73:*
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	3.5	111	2	A85866 hypothetical prote
2	9	3.2	733	2	578376 photosystem I P700
3	8	2.8	58	2	A58208 protamine I-1 - pa
4	8	2.8	143	2	G84168 hypothetical prote
5	8	2.8	197	2	E72374 hypothetical prote
6	8	2.8	220	2	AG3547 bicyclomycin resis
7	8	2.8	278	2	D83080 hypothetical prote
8	8	2.8	339	2	C71132 hypothetical prote
9	8	2.8	372	2	H70813 probable cysteine
10	8	2.8	379	2	F64300 formate dehydrogen
11	8	2.8	381	2	AH3041 conserved hypotet
12	8	2.8	387	2	D84885 hypothetical prote
13	8	2.8	397	2	D98244 hypothetical prote
14	8	2.8	422	1	A60503 sperm-binding glyco
15	8	2.8	443	2	T17220 hypothetical prote
16	8	2.8	465	2	AC0347 probable membrane
17	8	2.8	471	2	A75267 probable transpor
18	8	2.8	576	2	E64186 probable ATP-bind
19	8	2.8	586	2	A41125 gamma-glutamyltran
20	7	2.5	45	2	D58208 protamine I1-3 - p
21	7	2.5	50	2	S22582 protamine I - Sagu
22	7	2.5	58	2	S34045 S22582
23	7	2.5	86	2	F87604 F87604
24	7	2.5	102	2	F87993 F87993
25	7	2.5	115	2	PH1560 protein ZC334.3 [i
26	7	2.5	115	2	H83201 Ig heavy chain V r
27	7	2.5	118	1	IEEC58 conserved hypotet
28	7	2.5	118	2	AE1753 conserved hypotet
29	7	2.5	123	2	AH2707 conserved hypotet

ALIGNMENTS

30	7	2.5	125	2	T27519 hypothetical prote
31	7	2.5	131	2	I52290 interleukin-13 - r
32	7	2.5	131	2	E30552 T-cell activation
33	7	2.5	146	2	T37116 probable transpos
34	7	2.5	147	2	A71217 hypothetical prote
35	7	2.5	150	2	T08734 hypothetical prote
36	7	2.5	157	2	S31078 seed allergen RA5
37	7	2.5	157	2	T02664 allergen - rice
38	7	2.5	157	2	A75567 conserved hypotet
39	7	2.5	157	2	E75530 hypothetical prote
40	7	2.5	160	2	S59925 allergen RA5B prec
41	7	2.5	161	1	DNEC17 outer membrane pro
42	7	2.5	161	2	D90651 histone-like prote
43	7	2.5	161	2	D85502 hypothetical prote
44	7	2.5	162	2	T24937 hypothetical prote
45	7	2.5	162	2	T31173 hypothetical prote
46	7	2.5	164	2	S76920 hypothetical prote
47	7	2.5	170	2	S44789 D2007.4 protein -
48	7	2.5	174	2	D87638 beta-lactoglobulin
49	7	2.5	180	1	LGST
50	7	2.5	180	1	LGSH
51	7	2.5	181	2	B60738 insulin-like growth
52	7	2.5	187	2	G85343 phospholipase A2-I
53	7	2.5	206	2	S72567 hypothetical prote
54	7	2.5	230	2	AH0692 conserved hypotet
55	7	2.5	231	2	B64920 probable membrane
56	7	2.5	231	2	E90921 hypothetical prote
57	7	2.5	231	2	A85770 hypothetical prote
58	7	2.5	231	2	C86665 amino acid ABC tra
59	7	2.5	233	2	S60767 ribonuclease III -
60	7	2.5	233	2	A80273 H+-transporting AT
61	7	2.5	234	2	G85098 probable sodium-tr
62	7	2.5	235	2	I64174 probable cytochrom
63	7	2.5	236	2	A63387 hypothetical prote
64	7	2.5	238	2	H70866 conserved hypotet
65	7	2.5	240	2	G83208

RESULT 1
A85866 hypothetical protein Z3516 [imported] - Escherichia coli (strain O157:H7, substrain E
C:Species: Escherichia coli
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 14-Sep-2001
C:Accession: A85866
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May
Miller, L.; Grobeck, E.J.; Davis, N.W.; Lam, A.; Dimalanta, E.; Potamousis, K.; Apoda
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: A85866
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-111 <STO>
A:Cross-references: GB:AE005174; NID:q12516604; PIDN:AMG57389.1; GSPDB:GN00145; UWGP:
A:Experimental source: strain O157:H7, substrain EDL933
A:Genetics:
A:Gene: Z3516

Query Match 3.5% Score 10: DB 2: Length 111;
Best Local Similarity 100.0% Pred. NO. 0.073;
Matches 10: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGTALACGL 72
|||||
DB 40 LGTALACGL 49

RESULT 2
S78376 photosystem I P700 apoprotein A2 - Odontella sinensis chloroplast

C:Species: chloroplast *Odontella sinensis*
 C:Date: 17-Feb-1998 #sequence_revision 26-Feb-1998 #text_change 20-Jun-2000
 C:Accession: S78376
 R:Kowalik, K.V.; Stoebe, B.; Schaffran, I.; Kroch-Pancic, P.; Fretler, U.
 Plant Mol. Biol. Rep. 13, 336-342, 1995
 A:Title: The Chloroplast Genome of a chlorophyll a+c- containing Alga, *Odontella sinensis*
 A:Reference number: S78238
 A:Accession: S78376
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-733 <KOR>
 A:Cross-references: EMBL:267753; NID:q1185127; PIDN:CA91749.1; PID:q1185266
 A:Note: the nucleotide sequence was submitted to the EMBL Data Library, November 1995
 C:Genetics:
 A:Gene: psab
 A:Genome: chloroplast
 C:Superfamily: photosystem I P700 apoprotein
 C:Keywords: chloroplast; electron transfer; membrane protein; membrane-associated complex

Query Match 3.2%; Score 9; DB 2; Length 733;
 Best Local Similarity 100.0%; Pred. No. 3.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGALACLG 71
 |||||
 Db 333 LGALACLG 341

RESULT 3
 A58208
 C:Species: Chrysemys picta (painted turtle)
 C:Date: 08-Nov-1996 #sequence_revision 08-Nov-1996 #text_change 07-May-1999
 C:Accession: A58208
 R:Hunt, J.G.; Kaslinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp, D.
 J. Biol. Chem. 271, 23547-23557, 1996
 A:Title: Proteomes of reptiles.
 A:Reference number: A58208; MUID:96394458; PMID:8798564
 A:Accession: A58208
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-58 <HUN>
 C:Superfamily: sperm histone

Query Match 2.8%; Score 8; DB 2; Length 58;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 ORRGRRG 49
 |||||
 Db 35 ORRGRRG 42

RESULT 4
 G84168
 hypothetical protein Vng0080h [imported] - *Halobacterium* sp. NRC-1
 C:Species: *Halobacterium* sp. NRC-1
 C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001
 C:Accession: G84168
 R:Ng, W.V.; Kennedy, S.P.; Mahalas, G.G.; Bergquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leitner, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo
 Jung, K.H.; Alam, M.; Fretler, T.
 Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000
 A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ehardt, H.; Lowe, T.M.; Li
 A:Title: Genome sequence of *Halobacterium* species NRC-1.
 A:Reference number: AB4160; MUID:20504483; PMID:11016950
 A:Accession: G84168
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-143 <STO>
 A:Cross-references: GB:AE004437; NID:q10579733; PIDN:AG18715.1; GSPDB:GN00138
 C:Genetics:
 A:Gene: VNG0080H

Query Match 2.8%; Score 8; DB 2; Length 143;
 Best Local Similarity 100.0%; Pred. No. 7.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGGLALA 68
 |||||
 Db 55 LGGLALA 62

RESULT 5
 E72374
 hypothetical protein - *Thermotoga maritima* (strain MSB8)
 C:Species: *Thermotoga maritima*
 C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
 C:Accession: E72374
 R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic
 Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,
 C.M.
 Nature 399, 323-329, 1999
 A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome
 A:Reference number: A72200; MUID:99287316; PMID:10360571
 A:Accession: E72374
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-197 <ARN>
 A:Cross-references: GB:AE001724; GB:AE000512; NID:q4980966; PIDN:AAD35553.1; PID:q498
 C:Genetics:
 A:Experimental source: strain MSB8
 A:Gene: TM0469

Query Match 2.8%; Score 8; DB 2; Length 197;
 Best Local Similarity 100.0%; Pred. No. 9.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 SLDFEIS 9
 |||||
 Db 135 SLDFEIS 142

RESULT 6
 AG3547
 bicyclomycin resistance protein [imported] - *Brucella melitensis* (strain 16M)
 C:Species: *Brucella melitensis*
 C:Date: 01-Feb-2002 #sequence_revision 01-Feb-2002 #text_change 01-Feb-2002
 C:Accession: AG3547
 R:DelVecchio, V.G.; Kaparatil, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov
 .; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let
 Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002
 A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melit*
 A:Reference number: AD3252; PMID:11756688
 A:Accession: AG3547
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-220 <KUR>
 A:Cross-references: GB:AE008918; PIDN:AAL53546.1; PID:q17984455; GSPDB:GN00191
 A:Experimental source: strain 16M
 C:Genetics:
 A:Gene: BMEI10304
 A:Map position: II

Query Match 2.8%; Score 8; DB 2; Length 220;
 Best Local Similarity 100.0%; Pred. No. 11;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 54 ALLVPLAL 61
 |||||
 Db 132 ALLVPLAL 139

RESULT 7
 D83080
 hypothetical protein PA4521 [imported] - *Pseudomonas aeruginosa* (strain PA01)

C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: D83080
 R:Stover, C.K.; Pham, X.Q.; Ervin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lm,
 ; Lory, S.; Olson, M.V.
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic patho
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: D83080
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-278 <STO>
 A:Cross-references: GB:AE004866; GB:AE004091; NID:g9950760; PIDN:AG07909.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA4521

Query Match 2.8%; Score 8; DB 2; Length 278;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAVV 77
 |||||
 Db 47 LGLLAVV 54

RESULT 8
 C71132
 hypothetical protein PH0824 - *Pyrococcus horikoshii*
 C:Species: *Pyrococcus horikoshii*
 C:Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000
 C:Accession: C71132
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hatakeyama, Y.; Hino, Y.; Yamamoto, S.; Sekir
 M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kishida, N.; Oguchi
 DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
 A:Reference number: A71000; MUID:98344137; PMID:9679194
 A:Accession: C71132
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-339 <RAW>
 A:Cross-references: GB:AP000003; NID:g3236130; PIDN:BAA29917.1; PID:g3257234
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by Genbank
 C:Genetics:
 A:Gene: PH0824
 C:Superfamily: conserved hypothetical protein MTH900

Query Match 2.8%; Score 8; DB 2; Length 339;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAVV 77
 |||||
 Db 111 LGLLAVV 118

RESULT 9
 H70813
 probable cysteine synthase - *Mycobacterium tuberculosis* (strain H37RV)
 C:Species: *Mycobacterium tuberculosis*
 C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
 C:Accession: H70813
 R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S
 ; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.
 Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
 Nature 393, 537-544, 1998
 A:Authors: Seares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
 A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome
 A:Reference number: A70500; MUID:98295987; PMID:9634230
 A:Accession: H70813
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA
 A:Residues: 1-372 <COL>
 A:Cross-references: GB:AL022004; GB:AL123456; NID:g3261550; PIDN:CA17654.1; PID:g291
 A:Experimental source: strain H37RV
 C:Genetics:
 A:Gene: cysM3
 C:Superfamily: threonine dehydratase

Query Match 2.8%; Score 8; DB 2; Length 372;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGGLALA 68
 |||||
 Db 96 LGGLALA 103

RESULT 10
 E64300
 formate dehydrogenase (EC 1.2.1.2) beta chain - *Methanococcus jannaschii*
 C:Species: *Methanococcus jannaschii*
 C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000
 C:Accession: E64300
 R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak
 ; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek,
 rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.
 Science 273, 1058-1073, 1996
 A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese
 A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannasc*
 A:Reference number: A64300; MUID:96337999; PMID:8688087
 A:Accession: E64300
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-379 <BNL>
 A:Cross-references: GB:U67459; GB:L77117; NID:g2826236; PIDN:AAB97986.1; PID:g1590818
 C:Genetics:
 A:Map position: REV7250-6111
 C:Superfamily: formate dehydrogenase chain B; ferredoxin 2[4Fe-4S] homology
 C:Keywords: oxidoreductase
 F:273-348/Domain: ferredoxin 2[4Fe-4S] homology <FER>

Query Match 2.8%; Score 8; DB 2; Length 379;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 217 LVGVYAL 224
 |||||
 Db 35 LVGVYAL 42

RESULT 11
 AH3041
 conserved hypothetical protein Atu3948 [imported] - *Agrobacterium tumefaciens* (strain
 C:Species: *Agrobacterium tumefaciens*
 C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002
 C:Accession: AH3041
 R:Mood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavyn, T.; Levy, R.; Li, M.; McCl
 ; Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam
 ster, E.W.
 A:Title: The genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.
 A:Reference number: AB2577; PMID:11743193
 A:Accession: AH3041
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-381 <KUR>
 A:Cross-references: GB:AE008689; PIDN:AL44750.1; PID:g17742385; GSPDB:GN00187
 A:Experimental source: strain C58 (Dupont)
 C:Genetics:
 A:Gene: Atu3948
 A:Map position: linear chromosome

Query Match 2.8%; Score 8; DB 2; Length 381;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRPGSS 258
 |||||||
 DB 325 LALRPGSS 332

RESULT 12

D84885

hypothetical protein At2g45000 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 02-Feb-2001

C:Accession: D84885

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentto, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanden, S.E.; Umayam, L.; Tallon, L.

eus, D.; Nieman, M.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: AB4420; MUID:20083487; PMID:10617197

A:Accession: D84885

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-387 <STO>

A:Cross-references: GB:AE002093; NID:94895250; PIDN:AA032835.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g45000

A:Map position: 2

Query Match 2.8%; Score 8; DB 2; Length 387;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 BELVAED 100
 |||||||
 DB 370 BELVAED 377

RESULT 13

D98244

hypothetical protein AGR_L1808 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C:Species: Agrobacterium tumefaciens

C:Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C:Accession: D98244

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,

A.; Liu, F.; Wollem, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A:Reference number: A97359; PMID:11743194

A:Accession: D98244

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-397 <KUR>

A:Cross-references: GB:AE007870; PIDN:AAK89478.1; PID:q15159347; GSPDB:GN00170

C:Genetics:

A:Gene: AGR_L1808

A:Map position: linear chromosome

Query Match 2.8%; Score 8; DB 2; Length 397;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRPGSS 258
 |||||||
 DB 341 LALRPGSS 348

RESULT 14

A60503

sperm-binding glycoprotein ZP3 precursor - golden hamster

N:Alternate names: sperm receptor; zona pellucida glycoprotein ZP3

C:Species: Mesocricetus auratus (golden hamster)
 C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: A60503

R:Kinloch, R.A.; Ruiz-Seller, B.; Wasserman, P.M.

Dev. Biol. 142, 414-421, 1990

A:Title: Genomic organization and polypeptide primary structure of zona pellucida gly

A:Reference number: A60503; MUID:91078540; PMID:2257975

A:Accession: A60503

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-422 <KIN>

A:Cross-references: GB:M63629

A:Note: the authors translated the codon CAA for residue 251 as Glu, and AGC for rest

C:Comment: This sulfated glycoprotein in the zona pellucida of the oocyte is a recept

C:Superfamily: sperm-binding glycoprotein ZP3; ZP domain homology

C:Keywords: glycoprotein; oocyte

F:45-300/Domain: ZP domain homology <ZPH>

Query Match 2.8%; Score 8; DB 1; Length 422;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGLGLA 66
 |||||||
 DB 386 LALGLGLA 393

RESULT 15

T17220

hypothetical protein DKFZp566001.1 - human

C:Species: Homo sapiens (man)

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 02-Jun-2000

C:Accession: T17220

R:Blum, H.; Bauersachs, S.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, September 1999

A:Reference number: Z18725

A:Accession: T17220

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-443 <BLU>

A:Cross-references: EMBL:AL117414

A:Experimental source: fetal kidney; clone DKFZp566001

C:Genetics:

A:Note: DKFZp566001.1

C:Superfamily: gamma-glutamyltransferase

Query Match 2.8%; Score 8; DB 2; Length 443;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68
 |||||||
 DB 14 LGIGLALA 21

RESULT 16

AC0347

probable membrane protein yegb [imported] - Yersinia pestis (strain C092)

C:Species: Yersinia pestis

C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 17-May-2002

C:Accession: AC0347

R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlball, R.W.; Holden, M.T.G.; Prentice, M.

deno-Parraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.

11. M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barr

Nature 413, 523-527, 2001

A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.

A:Reference number: AB0001; MUID:21470413; PMID:11586360

A:Accession: AC0347

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-465 <KUR>

A:Cross-references: GB:AL590842; PIDN:CA092102.1; PID:q15980820; GSPDB:GN00175

C:Genetics:

A:Gene: yegb

C:Superfamily: multidrug-efflux transporter

Query Match 2.8%; Score 8; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVSL 79
DB 335 LLLAVVSL 342

RESULT 17

A:5267

probable transport protein - *Deinococcus radiodurans* (strain R1)

C:Species: *Deinococcus radiodurans*

C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: A75267

R.White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.; M

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.

A:Reference number: A75250; MID:20036896; PMID:10567266

A:Accession: A75267

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-471 <WHI>

A:Cross-references: GB:AE002079; GB:AE00513; MID:g6460315; PIDN:AAF12043.1; PID:g646032

A:Experimental source: strain R1

A:Gene: DR2502

A:Map position: 1

Query Match 2.8%; Score 8; DB 2; Length 471;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLAL 67
DB 366 ALGLGLAL 373

RESULT 18

E64186

probable ATP-binding transport protein H1156 - *Haemophilus influenzae* (strain Rd KW20)

C:Species: *Haemophilus influenzae*

C>Date: 18-Aug-1995 #sequence_revision 18-Aug-1995 #text_change 02-Feb-2001

C:Accession: E64186

R.Fleischmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kiehlavag, A

; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Meldman, J

, D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghagen, N.S.M.

Science 269, 496-512, 1995

A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter,

A:Title: Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd.

A:Reference number: A64000; MID:95350630; PMID:7542800

A:Accession: E64186

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-576 <TIGR>

A:Cross-references: GB:U32795; GB:I42023; MID:91574708; PIDN:AAC22811.1; PID:91574713; T

C:Superfamily: unassigned ATP-binding cassette proteins; ATP-binding cassette homology

C:Keywords: ATP; nucleotide binding; P-loop

F:355-550/Domain: ATP-binding cassette homology <ABC>

F:372-379/Region: nucleotide-binding motif A (P-loop)

Query Match 2.8%; Score 8; DB 2; Length 576;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 58 PLALGLGL 65
DB 159 PLALGLGL 166

RESULT 19

A41125

gamma-glutamyltransferase (EC 2.3.2.2) related protein - human

N/Alternate names: gamma-glutamyltransferase-like activity 1; GGT-REL

C:Species: *Homo sapiens* (man)

C>Date: 27-Mar-1992 #sequence_revision 27-Mar-1992 #text_change 18-Jun-1999

C:Accession: A41125

R.Heisterkamp, N.; Rajpert-De Meyts, E.; Uribe, L.; Forman, H.J.; Groffen, J.

Proc. Natl. Acad. Sci. U.S.A. 88, 6303-6307, 1991

A:Title: Identification of a human gamma-glutamyl cleaving enzyme related to, but dis

A:Reference number: A41125; MID:91296809; PMID:1676842

A:Accession: A41125

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-586 <HEI>

A:Cross-references: GB:M64099; MID:9183141; PIDN:AAA8503.1; PID:9183142

C:Genetics: GDB:GOTL1; GGT-REL

A:Gene: GDB:GOTL1; GGT-REL

A:Cross-references: GDB:13403

C:Superfamily: gamma-glutamyltransferase

C:Keywords: aminocyltransferase; glycoprotein; transmembrane protein

Query Match 2.8%; Score 8; DB 2; Length 586;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGGLALA 68
DB 14 LGGLALA 21

RESULT 20

D58208

protamine II-3 - painted turtle

C:Species: *Chrysemys picta* (painted turtle)

C>Date: 08-Nov-1996 #sequence_revision 08-Nov-1996 #text_change 07-May-1999

C:Accession: D58208

R.Hunt, J.G.; Kasinsky, H.E.; Elsey, R.M.; Wright, C.L.; Rice, P.; Bell, J.E.; Sharp,

J. Biol. Chem. 271, 23547-23557, 1996

A:Title: Protamines of reptiles.

A:Reference number: A58208; MID:96394458; PMID:8798564

A:Accession: D58208

A:Status: preliminary

A:Molecule type: protein

A:Residues: 1-45 <HUN>

C:Superfamily: sperm histone

Query Match 2.5%; Score 7; DB 2; Length 45;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49
DB 24 RRRGRG 30

RESULT 21

S22582

protamine 1 - *Saguinus imperator*

C:Species: *Saguinus imperator*

C>Date: 22-Nov-1993 #sequence_revision 12-Apr-1996 #text_change 21-Jul-2000

C:Accession: S22582

R.Queral, R.; Oliva, R.

Nucleic Acids Res. 19, 5786, 1991

A:Title: Protamine 1 gene sequence from the primate *Saguinus imperator* isolated with

A:Reference number: S22582; MID:92051332; PMID:1840669

A:Accession: S22582

A:Molecule type: DNA

A:Residues: 1-50 <QUE>

A:Cross-references: EMBL:X61678; MID:958405; PIDN:CAA43853.1; PID:94494091

A>Note: the authors translated the codon TAC for residue 43 as Thr

C:Genetics:

A:Introns: 37/1
 C:Superfamily: sperm histone
 C:Keywords: chromosomal protein; DNA binding; nucleus; spermatogenesis

Query Match 2.5%; Score 7; DB 2; Length 50;
 Best Local Similarity 100.0%; Pred. No. 28;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 42 QRRGR 48
 |||||
 Db 18 QRRGR 24

RESULT 22

S34045
 protamine - North American opossum
 C:Species: Didelphis virginiana, Didelphis marsupialis virginiana (North American opossum)
 C:Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 23-Jul-1999

C:Accession: S34045
 R:Winkler, R.J.; Nishikawa, S.; Connor, W.; Dixon, G.H.
 Eur. J. Biochem. 215, 63-72, 1993

A:Title: Characterization of a marsupial sperm protamine gene and its transcripts from t
 A:Reference number: S34045; MUID:93345500; PMID:8344286

A:Accession: S34045
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-58 <STO>
 A:Cross-references: EMBL:X74044; NID:g407062; PIDN:CAA52193.1; PID:g407063
 C:Superfamily: sperm histone
 C:Keywords: DNA binding; nucleus

Query Match 2.5%; Score 7; DB 2; Length 58;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49
 |||||
 Db 35 RRRGRG 41

RESULT 23

F87604
 hypothetical protein CC2870 [imported] - Caulobacter crescentus
 C:Species: Caulobacter crescentus
 C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001

C:Accession: F87604
 R:Nierman, W.C.; Feldlyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001

A:Title: Complete genome sequence of *Caulobacter crescentus*.
 A:Reference number: AB7249; MUID:21173698; PMID:11259647
 A:Accession: F87604
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-86 <STO>
 A:Cross-references: GB:AE005673; NID:g13424486; PIDN:AAK24834.1; GSPDB:GN00148

C:Genetics:
 A:Gene: CC2870

Query Match 2.5%; Score 7; DB 2; Length 86;
 Best Local Similarity 100.0%; Pred. No. 45;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LALGLGL 65
 |||||
 Db 26 LALGLGL 32

RESULT 24

F87993
 protein ZC334.3 [imported] - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*

C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
 C:Accession: F87993
 R:Anonymous, The C. elegans Sequencing Consortium.
 Science 282, 2012-2018, 1998

A:Title: Genome sequence of the nematode *C. elegans*: a platform for investigating bio
 A:Reference number: AJ5000; MUID:99069613; PMID:9851916
 A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_

A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999;
 A:Accession: F87993
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-102 <STO>
 A:Cross-references: GB:chr_I; PIDN:CA04964.1; PID:g3881432; GSPDB:GN00019; CESP:ZC33

C:Genetics:
 A:Gene: ZC334.3
 A:Map position: 1

Query Match 2.5%; Score 7; DB 2; Length 102;
 Best Local Similarity 100.0%; Pred. No. 52;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 55 LVPPLAL 61
 |||||
 Db 33 LVPPLAL 39

RESULT 25

PH1560
 Ig heavy chain V region (clone VH32) - human (fragment)

C:Species: Homo sapiens (man)
 C:Date: 05-Aug-1994 #sequence_revision 05-Aug-1994 #text_change 21-Jan-2000
 C:Accession: PH1560

R:Rassenti, L.Z.; Kipps, T.J.
 J. Exp. Med. 177, 1039-1046, 1993

A:Title: Lack of extensive mutations in the VH5 genes used in common B cell chronic 1
 A:Reference number: PH157; MUID:93210459; PMID:7681468

A:Accession: PH1560
 A:Molecule type: DNA
 A:Residues: 1-115 <RAS>
 C:Superfamily: immunoglobulin V region; immunoglobulin homology
 C:Keywords: heterotrimer; immunoglobulin
 F:32-115/Domain: immunoglobulin homology <IMM>

Query Match 2.5%; Score 7; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLLAV 76
 |||||
 Db 5 LGLLLAV 11

RESULT 26

H83201

conserved hypothetical protein PA3557 [imported] - *Pseudomonas aeruginosa* (strain PAO
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: H83201
 R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;
 Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000

A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pa
 A:Reference number: AB2950; MUID:20437337; PMID:10964043

A:Accession: H83201

A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-115 <STO>

A:Cross-references: GB:AE004776; GB:AE004091; NID:g9949701; PIDN:AA606945.1; GSPDB:GN
 A:Experimental source: strain PAO1

C:Genetics:
 A:Gene: PA3557

Query Match 2.5%; Score 7; DB 2; Length 115;
 Best Local Similarity 100.0%; Pred. No. 57;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 59 LALGIGL 65
 |||||
 Db 49 LALGIGL 55

RESULT 27

IECSB
 hypothetical protein, 12K - Escherichia coli insertion sequence ISS

C:Species: Escherichia coli
 C:Date: 18-Dec-1981 #sequence_revision 18-Dec-1981 #text_change 10-Sep-1999

C:Accession: B91483; C03582; A04466

R:Schoner, B.; Kahn, M.

Gene 14, 165-174, 1981

A:Title: The nucleotide sequence of ISS from Escherichia coli.

A:Reference number: A91483; MUID:82028653; PMID:6269959

A:Accession: B91483

A:Molecule type: DNA

A:Residues: 1-118 <SCH>

R:Engler, J.A.; van Bree, M.P.

Gene 14, 155-163, 1981

A:Title: The nucleotide sequence and protein-coding capability of the transposable element

A:Reference number: A03582; MUID:82028652; PMID:6269958

A:Accession: C03582

A:Molecule type: DNA

A:Residues: 1-118 <ENG>

C:Genetics:

A:Mobile element: Insertion sequence ISS

C:Superfamily: Escherichia coli hypothetical protein o263

Query Match 2.5%; Score 7; DB 1; Length 118;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 23 DGAVRQ 29
 |||||
 Db 107 DGAVRQ 113

RESULT 28

AE1753

Orf1 (bacteriophage b1285) homolog lin2570 [imported] - Listeria innocua (strain Clp1)

C:Species: Listeria innocua

C:Date: 27-Nov-2001 #sequence_revision 27-Nov-2001 #text_change 27-Nov-2001

C:Accession: AE1753

R:Glaser, P.; Frangoul, L.; Buchrieser, C.; Amend, A.; Baquero, F.; Berche, P.; Bloeker, D.; Dominguez-Bernal, G.; Duchaud, E.; Durand, L.; Dussurget, O.; Entian, K.D.; Fsihl, H.

Science 294, 849-852, 2001

A:Authors: Kieft, J.; Kuhn, M.; Kunst, F.; Kurapkat, G.; Madheno, E.; Maitournam, A.; Maok, C.; Schluter, T.; Simoes, N.; Tlerraz, A.; Vazquez-Boland, J.A.; Voss, H.; Wehland, A.

A:Title: Comparative genomics of Listeria species.

A:Reference number: AB1077; MUID:21537279; PMID:11679669

A:Accession: AE1753

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-118 <GLA>

A:Cross-references: GB:AL592022; PIDN:CAC97797.1; PID:916415092; GSPDB:GN00178

A:Experimental source: strain Clp11262

C:Genetics:

A:Gene: lin2570

Query Match 2.5%; Score 7; DB 2; Length 118;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 2 SLDFEI 8
 |||||
 Db 61 SLDFEI 67

RESULT 29

AH2707

conserved hypothetical protein Atu1065 [imported] - Agrobacterium tumefaciens (strain

C:Species: Agrobacterium tumefaciens

C:Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AH2707

R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moe

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutlyavin, T.; Levy, R.; Li, M.; McCl

ard, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AH2707

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-123 <KUR>

A:Cross-references: GB:AE008688; PIDN:AL42078.1; PID:917739458; GSPDB:GN00186

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Map position: circular chromosome

A:Gene: Atu1065

Query Match 2.5%; Score 7; DB 2; Length 123;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 12 RLPLPRS 18
 |||||
 Db 50 RLPLPRS 56

RESULT 30

T27519

hypothetical protein ZC334.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C:Accession: T27519

R:McLay, K.

submitted to the EMBL Data Library, November 1996

A:Reference number: Z20381

A:Accession: T27519

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-125 <WIL>

A:Cross-references: EMBL:Z82082; PIDN:CA804964.2; GSPDB:GN00019; CESP:ZC334.3

A:Experimental source: clone ZC334

C:Genetics:

A:Gene: CESP:ZC334.3

A:Map position: 1

A:introns: 99/1

Query Match 2.5%; Score 7; DB 2; Length 125;
 Best Local Similarity 100.0%; Pred. No. 61;
 Matches 7; Conservative 0; Mismatches 0; Indels 0;

OY 55 LVLPLAL 61
 |||||
 Db 33 LVLPLAL 39

RESULT 31

I52290

Interleukin-13 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 16-Jul-1999

C:Accession: I52290

R:Lakkis, F.G.; Cruet, E.N.

Biochem. Biophys. Res. Commun. 197, 612-618, 1993

A:Title: Cloning of rat interleukin-13 (IL-13) cDNA and analysis of IL-13 gene expres

A:Reference number: I52290; MUID:94092138; PMID:7916015
 A:Accession: I52290
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-131 <RES>
 A:Cross-references: GB:U26913; NID:9438875; PIDN:AAAI6478.1; PID:9438876
 C:Genetics:
 A:Gene: IL-13
 C:Superfamily: Interleukin-13

Query Match 2.5%; Score 7; DB 2; Length 131;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACLG 71
 |||||
 DB 9 LALACLG 15

RESULT 32

E30552

T-cell activation protein P600 precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 28-Aug-1989 #sequence_revision 28-Aug-1989 #text_change 16-Jul-1999

C:Accession: E30552

R:Brown, K.D.; Zurawski, S.M.; Mosmann, T.R.; Zurawski, G.

J. Immunol. 142, 679-687, 1989

A:Title: A family of small inducible proteins secreted by leukocytes are members of a new class of various activation processes.

A:Reference number: A30552; MUID:89093958; PMID:2521353

A:Accession: E30552

A:Molecule type: mRNA

A:Residues: 1-131 <BRO>

A:Cross-references: GB:M23504; NID:9533246; PIDN:AAA40149.1; PID:9533247

C:Superfamily: Interleukin-13

Query Match 2.5%; Score 7; DB 2; Length 131;
 Best Local Similarity 100.0%; Pred. No. 64;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACLG 71
 |||||
 DB 9 LALACLG 15

RESULT 33

T37116

probable transposase, truncated [imported] - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 08-Sep-2000 #sequence_revision 08-Sep-2000 #text_change 15-Sep-2000

C:Accession: T37116

R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21588

A:Accession: T37116

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-146 <SAU>

A:Cross-references: EMBL:AL109950; PIDN:GAB52967.1; GSPDB:GN00070; SCOEDB:SCJ4.33c

A:Experimental source: strain A312

C:Genetics:

A:Gene: SCOEDB:SCJ4.33c

C:Superfamily: Synchocystis transposase s111710

Query Match 2.5%; Score 7; DB 2; Length 146;
 Best Local Similarity 100.0%; Pred. No. 70;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAV 76
 |||||
 DB 102 LGLLAV 108

RESULT 34

A71217

hypothetical protein PH2001 - Pyrococcus horikoshii

C:Species: Pyrococcus horikoshii

C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 20-Jun-2000

C:Accession: A71217

R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hino, Y.; Yamamoto, S.; Se

M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kusuda, N.; Ogu

DNA Res. 5, 55-76, 1998

A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic

A:Reference number: A71000; MUID:98344137; PMID:9679194

A:Accession: A71217

A:Status: nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-147 <KAW>

A:Cross-references: GB:AP000007; GB:AP000001; NID:g3236134; NID:g3236128; PID:g325844

A:Experimental source: strain OT3

A:Note: this accession replaces an interim accession for a sequence replaced by GenBank

C:Genetics:

A:Gene: PH2001

C:Superfamily: Pyrococcus horikoshii hypothetical protein PH2001

Query Match 2.5%; Score 7; DB 2; Length 147;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 67 LACIGLL 73
 |||||
 DB 45 LACIGLL 51

RESULT 35

T08734

hypothetical protein DKFZP566F0546.1 - human (fragment)

C:Species: Homo sapiens (man)

C>Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 13-Aug-1999

C:Accession: T08734

R:Ottewill, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, May 1999

A:Reference number: Z16474

A:Accession: T08734

A:Molecule type: mRNA

A:Residues: 1-150 <OTT>

A:Cross-references: EMBL:AL050075

A:Experimental source: fetal kidney; clone DKFZP566F0546

C:Genetics:

A:Note: DKFZP566F0546.1

Query Match 2.5%; Score 7; DB 2; Length 150;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 234 AASSLGP 240
 |||||
 DB 32 AASSLGP 38

RESULT 36

S31078

seed allergen RA5 - rice

C:Species: Oryza sativa (rice)

C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 20-Jun-2000

C:Accession: S31078

R:Adachi, T.; Izumi, H.; Yamada, T.; Tanaka, K.; Takeuchi, S.; Nakamura, R.; Matsuda,

Plant Mol. Biol. 21, 239-248, 1993

A:Title: Gene structure and expression of rice seed allergenic proteins belonging to

A:Reference number: S31078; MUID:93144699; PMID:7678765

A:Accession: S31078

A:Molecule type: mRNA

A:Residues: 1-157 <ADA>

A:Cross-references: EMBL:D11430; NID:g218196; PIDN:BA01996.1; PID:g218197

C:Superfamily: wheat alpha-amylase inhibitor

C:Keywords: seed

Query Match 2.5%; Score 7; DB 2; Length 157;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVS 78
 |||||
 Db 11 LLLAVS 17

RESULT 37

T02664

allergen - rice

C:Species: Oryza sativa (rice)

C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 16-Jul-1999

C:Accession: T02664

R:Yun, C.H.; Park, J.H.; Eun, M.Y.

A:Submitted to the EMBL Data Library, January 1998

A:Description: Nucleotide sequence of rice allergenic protein.

A:Reference number: 214691

A:Accession: T02664

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-157 <YUN>

A:Cross-references: EMBL:AF042200; NID:q2827315; PIDN:AAF99797.1; PID:q2827316

A:Experimental source: strain Nipponbare

C:Superfamily: wheat alpha-amylase inhibitor

Query Match 2.5%; Score 7; DB 2; Length 157;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVS 78
 |||||
 Db 11 LLLAVS 17

RESULT 38

A75567

conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 17-Mar-2000

C:Accession: A75567

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: A75567

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 <WHI>

A:Cross-references: GB:AE001867; GB:AE000513; NID:q6457693; PIDN:AAF09626.1; PID:q6457693

C:Genetics:

A:Experimental source: strain R1

A:Gene: DR0033

A:Map position: 1

Query Match 2.5%; Score 7; DB 2; Length 157;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 261 IRTLPWA 267
 |||||
 Db 22 IRTLPWA 28

RESULT 39

E75530

hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000

C:Accession: E75530

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: A75250; MUID:20036896; PMID:10567266

A:Accession: E75530

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 <WHI>

A:Cross-references: GB:AE001895; GB:AE000513; NID:q6458024; PIDN:AAF09935.1; PID:q645

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR0352

A:Map position: 1

Query Match 2.5%; Score 7; DB 2; Length 157;
 Best Local Similarity 100.0%; Pred. No. 75;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRGRRG 49
 |||||
 Db 144 RRGRRG 150

RESULT 40

S59925

allergen RASB precursor - rice

C:Species: Oryza sativa (rice)

C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 20-Jun-2000

C:Accession: S59925

R:Alvarez, A.M.; Adachi, T.; Nakase, M.; Aoki, N.; Nakamura, R.; Matsuda, T.

Biochim. Biophys. Acta 1251, 201-204, 1995

A:Title: Classification of rice allergenic protein cDNAs belonging to the alpha-amylase

A:Reference number: S59922; MUID:95399441; PMID:7669811

A:Accession: S59925

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-160 <ALV>

A:Cross-references: EMBL:D42142; NID:q1398917; PIDN:BAA07713.1; PID:q1398918

C:Superfamily: wheat alpha-amylase inhibitor

Query Match 2.5%; Score 7; DB 2; Length 160;
 Best Local Similarity 100.0%; Pred. No. 76;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVS 78
 |||||
 Db 11 LLLAVS 17

Search completed: March 31, 2003, 14:16:42
 Job time : 47 secs

THIS PAGE BLANK (USPTO)

Brue
RS 5/15/03
Page 1

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 11:55:46 ; Search time 22 Seconds
(without alignments)
535.422 Million cell updates/sec

Title: US-09-245-198A-4
Perfect score: 284
Sequence: 1 MSLLDFEISARRLPRLPSLG.....PMAHLKAAPRLTYFGLFQVH 284

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 65 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	249	87.7	249	1	043508 Homo sapien
2	32	11.3	225	1	TN12_HUMAN
3	10	3.5	111	1	YFBW_ECOLI
4	9	3.2	733	1	PSAB_ODOSI
5	9	3.2	734	1	PSAB_CVACA
6	8	2.8	179	1	ADHS_GLUOX
7	8	2.8	220	1	Y304_BRUME
8	8	2.8	317	1	MSHR_PANTR
9	8	2.8	379	1	FDHB_MERJA
10	8	2.8	422	1	2P3_MESAV
11	8	2.8	576	1	CYDC_HUMAN
12	8	2.8	586	1	GGT5_HUMAN
13	7	2.5	49	1	HSP1_SAGIM
14	7	2.5	57	1	HSP1_DIDMA
15	7	2.5	115	1	A62F_DROME
16	7	2.5	118	1	Y151_ECOLI
17	7	2.5	131	1	IL13_MOUSE
18	7	2.5	131	1	IL13_MOUSE
19	7	2.5	147	1	YK01_PYRHO
20	7	2.5	150	1	TNFC_PIG
21	7	2.5	157	1	RA05_ORYSA
22	7	2.5	161	1	HLPA_ECOLI
23	7	2.5	170	1	YLMA_CAEEL
24	7	2.5	180	1	LACB_BUBBU
25	7	2.5	180	1	LACB_CAPI
26	7	2.5	180	1	LACB_SHEEP
27	7	2.5	230	1	RNFE_SALTY
28	7	2.5	231	1	RNFE_ECO57
29	7	2.5	231	1	RNFE_ECOLI
30	7	2.5	233	1	RNFE_ECOLI
31	7	2.5	233	1	RNFE_COXBU
32	7	2.5	235	1	RNFE_YERPE
33	7	2.5	239	1	TN14_MOUSE

ALIGNMENTS

34	7	2.5	240	1	RNFE_PSEAE	09yhb5 pseudomonas
35	7	2.5	243	1	MOEB_HAEIN	P45211 haemophilus
36	7	2.5	244	1	TNFC_HUMAN	006643 homo sapien
37	7	2.5	249	1	MOEB_ECOLI	P12282 escherichia
38	7	2.5	249	1	MOEB_SALTY	056057 salmonella
39	7	2.5	257	1	KDTX_SERMA	054435 serattia ma
40	7	2.5	310	1	TNFC_MARMO	09jml0 marmota mon
41	7	2.5	316	1	ISPH_XYLFA	09pas9 xyliella fas
42	7	2.5	329	1	ODPB_BACSU	P21882 bacillus su
43	7	2.5	329	1	SRA6_CAEEL	009208 caenorhabdi
44	7	2.5	335	1	LAG2_RAT	035777 rattus norv
45	7	2.5	344	1	LEU3_THERAO	P24098 thermus aqu
46	7	2.5	357	1	G6PT_MOUSE	P35576 mus musculu
47	7	2.5	357	1	G6PT_RAT	P43428 rattus norv
48	7	2.5	358	1	PONT_RABIT	P27170 oryctolagus
49	7	2.5	365	1	MOEB_THERH	060019 thermus the
50	7	2.5	394	1	BENE_ACICA	P07775 acinetobact
51	7	2.5	396	1	DHH1_XENLA	091610 xenopus lae
52	7	2.5	398	1	DHH2_XENLA	091611 xenopus lae
53	7	2.5	402	1	SELP_BOVIN	P49907 bos taurus
54	7	2.5	412	1	PKRP_ALCEU	P50320 alcaligenes
55	7	2.5	413	1	PKRC_ALCEU	P50319 alcaligenes
56	7	2.5	416	1	NH59_CAEEL	09txj1 caenorhabdi
57	7	2.5	418	1	CP16_RAT	P09006 rattus norv
58	7	2.5	419	1	ENO_PYRAE	08zye7 pyrobaculum
59	7	2.5	423	1	YJ54_YEAST	P47130 saccharomyc
60	7	2.5	424	1	TP3_MOUSE	P10761 mus musculu
61	7	2.5	428	1	SYH_CHLMU	09pj19 ciliandia m
62	7	2.5	455	1	PHR_STRGR	P12768 streptomyce
63	7	2.5	461	1	PUCG_RHOCA	P23462 rhodobacter
64	7	2.5	461	1	Y608_HAEIN	057486 haemophilus
65	7	2.5	464	1	SOXB_MOUSE	004886 mus musculu

RESULT 1

ID	TN12_HUMAN	STANDARD:	PRT:	249 AA.
AC	043508; Q8WU27;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (APO3 ligand).			
GN	TNFSF12 OR APO3L OR DR3LG.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A. AND N-TERMINUS OF SOLUBLE FORM.			
RC	TISSUE=Tonsil, and Fetal liver;			
RX	MEDLINE=98070415; PubMed=9405449;			
RA	Chicheportliche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H., Heeston C., Garcia I., Browning J.L., a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis.			
RT	J. Biol. Chem. 272:32401-32410(1997).			
RL	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fetal kidney;			
RX	MEDLINE=98228355; PubMed=9560343;			
RA	Marsters S.A., Sheridan J.P., Pilti R.M., Brush J., Goddard A., Ashkenazi A.;			
RT	"Identification of a ligand for the death-domain-containing receptor APO3".			
RL	Curr. Biol. 8:525-528(1998).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Tonsil;			
RA	Strausberg R.;			

```

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RP FUNCTION.
RA Pubmed=10085077;
RA Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;
RT "TNF $\alpha$  induces angiogenesis and proliferation of endothelial cells.";
RL J. Biol. Chem. 274:8455-8459(1999).
CC -1- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/AP03. Weak
CC inducer of apoptosis in some cell types. Mediates NF-kappa
CC activation. May promote angiogenesis and the proliferation of
CC endothelial cells.
CC -1- SUBUNIT: HOMODIMER (POTENTIAL).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted.
CC -1- TISSUE SPECIFICITY: Highly expressed in adult heart, pancreas,
CC skeletal muscle, brain, colon, small intestine, lung, ovary,
CC prostate, spleen, lymph node, appendix and peripheral blood
CC lymphocytes. Low expression in kidney, testis, liver, placenta,
CC thymus and bone marrow. Also detected in fetal kidney, liver,
CC lung and brain.
CC -1- PTM: The soluble form derives from the membrane form
CC by proteolytic processing.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -1- CAUTION: Ref.3 sequence differs from that shown due to a
CC frameshift in position 125.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF030099; AAC51923.1; -
DR EMBL: AF055872; AAC39724.1; -
DR EMBL: BC019047; AAH19047.1; ALT_FRAME.
DR Genes: HGNC:11927; TNFSF12.
DR MIM: 602695; -
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_2; 1. FALSE_NEG.
DR PROSITE: PS50049; TNF_2; 1.
KM Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 249
FT FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT FT MEMBER 12, MEMBRANE FORM.
FT FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT CHAIN 94 249
FT FT MEMBER 12, SECRETED FORM.
FT FT CYTOPLASMIC (POTENTIAL).
FT FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT TRANSMEM 22 42
FT FT (POTENTIAL).
FT DOMAIN 43 249
FT SITE 93 94
FT CARBOHYD 139 139
FT SEQUENCE 249 AA: 27216 MW: E66084361C28B8A CRC64:
SO QUERY MATCH 87.7%; Score 249; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 3.2e-231;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DB 181 LKVDGVLALRCLEEFSAATASSLGLPOLRLCQVSGLLALRPGSSLRITRLPMHLKAPFL 240
OY 276 TYEGLFQVH 284
DB 241 TYEGLFQVH 249
RESULT 2
TNF2_MOUSE STANDARD: PRT: 225 AA.
AC 054907: 09CTP2:
ID TNF2_MOUSE
DT 15-JUN-2002 (rel. 41, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak
DE inducer of apoptosis) (TNF $\alpha$ ) (Fragment).
GN TNFSF12
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Peritoneal macrophage;
RX MEDLINE=98070415; Pubmed=9405449;
RA Chicheportliche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,
RA Hession C., Garcia I., Browning J.L.;
RT "TNF $\alpha$ , a new secreted ligand in the tumor necrosis factor family that
RT weakly induces apoptosis.";
RL J. Biol. Chem. 272:32401-32410(1997).
RN [2]
RP SEQUENCE OF 83-225 FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Retina;
RX MEDLINE=21085660; Pubmed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadoya K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochila H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Queckenbush J.,
RA Schimi L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baladrelli R., Barsch G.,
RA Blake J., Boilelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Williams L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohlsuki S.,
RA Hayashida Y.;
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
CC -1- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/AP03. Weak
CC inducer of apoptosis in some cell types. Promotes angiogenesis and
CC the proliferation of endothelial cells. Mediates NF-kappa
CC activation (By similarity).
CC -1- SUBUNIT: Homotrimer (Potential).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).
CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.
CC -1- PTM: The soluble form is produced from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```



```

FT TRANSMEM 174 198 III (POTENTIAL).
FT TRANSMEM 272 290 IV (POTENTIAL).
FT TRANSMEM 329 352 V (POTENTIAL).
FT TRANSMEM 368 394 VI (POTENTIAL).
FT TRANSMEM 416 438 VII (POTENTIAL).
FT TRANSMEM 516 534 VIII (POTENTIAL).
FT TRANSMEM 574 595 IX (POTENTIAL).
FT TRANSMEM 642 664 X (POTENTIAL).
FT TRANSMEM 706 726 XI (POTENTIAL).
FT METAL 558 558 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT METAL 567 567 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT BINDING 653 653 P700 SPECIAL PAIR CHLOROPHYLL AXIAL
FT BINDING 661 661 LIGAND (BY SIMILARITY).
FT BINDING 669 669 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 670 670 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 699 699 A1 PHYLOQUINONE (BY SIMILARITY).
FT BINDING 699 699 A1 PHYLOQUINONE (BY SIMILARITY).
SQ SEQUENCE 733 AA; 82103 MW; 13439AF1E41BBF7 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 1; Length 733;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGIALACLG 71
DB 333 LGIALACLG 341

RESULT 5
PSAB_CYACA STANDARD; PRT: 734 AA.
O9TID6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Photosystem I P700 chlorophyll A apoprotein A2 (Psab) (PST-B).
GN PSAB.
OS Cyanidium caldarium.
OC Chloroplast.
OC Eukaryota; Rhodophyta; Bangiophyceae; Porphyridiales; Porphyridiaceae;
OC Cyanidium.
OX NCBI_TaxID=2771;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-RK-1;
RX MEDLINE=20496959; PubMed=11040290;
RA Gloeckner G., Rosenthal A., Valentin K.-U.;
RT "The structure and gene repertoire of an ancient red algal plastid genome.";
RT J. Mol. Evol. 51:382-390(2000).
RL J. Mol. Evol. 51:382-390(2000).
CC -1- FUNCTION: Psaa and psab bind P700, the primary electron donor of photosystem I (PSI), as well as the electron acceptors A0, A1, and FX. PSI functions as a plastocyanin/cytochrome c6-ferredoxin oxidoreductase.
CC -1- COFACTOR: P700 is a chlorophyll A dimer, A0 is chlorophyll A, A1 is a phytylquinone and FX is a 4Fe-4S iron-sulfur center.
CC -1- SUBUNIT: A psaa/b heterodimer binds the P700 chlorophyll special pair and subsequent electron acceptors. The PSI reaction center of higher plants and algae is composed of one at least 11 subunits.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Chloroplast thylakoid membrane.
CC -1- SIMILARITY: BELONGS TO THE PSAA/PSAB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
CC EMBL: AF023186; AAF12881.1; -
CC HSSP: P25897; IJBO.

```

```

DR InterPro: IPR001280; PSI_Psaa/B.
DR Pfam: PF00223; psaa-psab; 1.
DR PRINTS: PR00257; PHOTOSYSAB.
DR PROSITE: PS00419; PHOTOSYSTEM_I_PSAAB; 1.
KM Photosynthesis; Photosystem I; Transport; Electron transport;
KM Chloroplast; Thylakoid; Transmembrane; Metal-binding; Iron;
KW Iron-sulfur; 4Fe-4S; Chlorophyll.
FT TRANSMEM 46 69 I (POTENTIAL).
FT TRANSMEM 135 158 II (POTENTIAL).
FT TRANSMEM 175 199 III (POTENTIAL).
FT TRANSMEM 273 291 IV (POTENTIAL).
FT TRANSMEM 330 353 V (POTENTIAL).
FT TRANSMEM 369 395 VI (POTENTIAL).
FT TRANSMEM 417 439 VII (POTENTIAL).
FT TRANSMEM 517 535 VIII (POTENTIAL).
FT TRANSMEM 575 596 IX (POTENTIAL).
FT TRANSMEM 643 665 X (POTENTIAL).
FT TRANSMEM 707 727 XI (POTENTIAL).
FT METAL 559 559 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT METAL 568 568 IRON-SULFUR FX (4FE-4S) (BY SIMILARITY).
FT BINDING 654 654 P700 SPECIAL PAIR CHLOROPHYLL AXIAL
FT BINDING 662 662 LIGAND (BY SIMILARITY).
FT BINDING 670 670 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 671 671 A0 CHLOROPHYLL (BY SIMILARITY).
FT BINDING 700 700 A1 PHYLOQUINONE (BY SIMILARITY).
FT BINDING 700 700 A1 PHYLOQUINONE (BY SIMILARITY).
SQ SEQUENCE 734 AA; 82359 MW; 4496AA2AE59CA9B9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 1; Length 734;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGIALACLG 71
DB 334 LGIALACLG 342

```

```

RESULT 6
ADHS_GLUOX STANDARD; PRT: 179 AA.
AC 005544;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Alcohol dehydrogenase 15 kDa subunit precursor (G3-ADH subunit III).
GN ADHS.
OS Gluconobacter oxydans (Gluconobacter suboxydans).
OC Bacteria; Proteobacteria; alpha subdivision; Acetobacteraceae;
OC Gluconobacter.
OX NCBI_TaxID=442;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 26-40.
RC STRAIN-IFO 12528;
RX MEDLINE=97208225; PubMed=9055427;
RA Kondo K., Horiuchi S.;
RT "Characterization of the genes encoding the three-component membrane-bound alcohol dehydrogenase from Gluconobacter suboxydans and their RT expression in Acetobacter pasteurianus.";
RL Appl. Environ. Microbiol. 63:1131-1138(1997).
CC -1- FUNCTION: NOT ESSENTIAL FOR ALCOHOL DEHYDROGENASE ACTIVITY.
CC -1- SUBUNIT: HETEROTRIMER (DEHYDROGENASE, CYTOCHROME AND PROTEIN ADHS), THAT FORMS THE ALCOHOL DEHYDROGENASE COMPLEX.
CC -1- SUBCELLULAR LOCATION: MEMBRANE-BOUND, FACING THE PERIPLASMIC SPACE (POTENTIAL).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).
CC -----
CC

```

DR EMBL; D86440; BAA19756.1; -
 KW Membrane: Periplasmic; Signal.
 FT SIGNAL 1 24 POTENTIAL.
 FT CHAIN 25 179 ALCOHOL DEHYDROGENASE 15 KDA SUBUNIT.
 FT MOD_RES 25 25 PYRROLIDONE CARBOXYLIC ACID.
 SQ SEQUENCE 179 AA; 19943 MW; P6AF24365B3C66 CRC64;

Query Match 2.8%; Score 8; DB 1; Length 179;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 LALGIGLA 66
 |||||
 DB 11 LALGIGLA 18

RESULT 7
 Y304_BRUME STANDARD; PRT; 220 AA.
 AC Q8YD73;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Hypothetical protein BMEI10304.
 GN BMEI10304.
 OS Brucella melitensis.
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 CC Brucellaceae; Brucella.
 OX NCBI_TaxID=29459;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=16M / ATCC 23456 / Biotype 1;
 RX MEDLINE=20020109; PubMed=11756688;
 RA Delvecchio V.G., Kapratel V., Redkar R.J., Patra G., Mujer C., Los T.,
 RA Ivanova N., Anderson I., Bhattacharya A., Lykidis A., Reznik G.,
 RA Jaelkovi L., Larsen N., D'Souza M., Bernal A., Mazur M., Goldsman E.,
 RA Seikov E., Elzer P.H., Hagius S., O'Callaghan D., Letesson J.-D.,
 RA Haselkorn R., Kyriades N., Overbeek R.;
 RT "The genome sequence of the facultative intracellular pathogen
 RT Brucella melitensis.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:443-448(2002).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
 CC -1- SIMILARITY: BELONGS TO THE DPF0191 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AE009669; AAL53546.1; -
 KW Hypothetical protein; Transmembrane; Complete proteome.
 FT TRANSMEM 20 39 POTENTIAL.
 FT TRANSMEM 54 72 POTENTIAL.
 FT TRANSMEM 85 104 POTENTIAL.
 FT TRANSMEM 124 146 POTENTIAL.
 FT TRANSMEM 153 175 POTENTIAL.
 FT TRANSMEM 179 198 POTENTIAL.
 SQ SEQUENCE 220 AA; 24815 MW; 182C0244743B17A CRC64;

Query Match 2.8%; Score 8; DB 1; Length 220;
 Best Local Similarity 100.0%; Pred. No. 4.9;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 54 ALVPLAL 61
 |||||
 DB 132 ALVPLAL 139

RESULT 8
 MSHR_PANTR

ID MSHR_PANTR STANDARD; PRT; 317 AA.
 AC Q9URK4;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Melanocyte stimulating hormone receptor (MSH-R) (Melanotropin
 DE receptor) (Melanocortin-1 receptor) (MCL-R).
 GN MCLR.
 OS Pan troglodytes (Chimpanzee).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eulheria; Primates; Catarrhini; Homiidae; Pan.
 OX NCBI_TaxID=9598;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Rees J.L., Harding R.M., Healy E., Jackson I.J., Ray A.J., Ellis N.S.,
 RA Flanagan N., Todd C., Dixon C., Matthews J.N., Sajantila A.,
 RA Birch-Machin M.A.;
 RT "Chimpanzee melanocortin 1 sequence.";
 RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR MSH (ALPHA, BETA AND GAMMA) AND ACTH. THE
 CC ACTIVITY OF THIS RECEPTOR IS MEDIATED BY G PROTEINS WHICH ACTIVATE
 CC ADENYLATE CYCLASE (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ245705; CAB53398.1; -
 DR InterPro: IPR000276; GPCR_Rhodpsn.
 DR Pfam: PF00001; 7tm_1; 1.
 DR PRINTS: PR00237; GPCRHOPOPSN.
 DR PROSITE: PS00237; G-PROTEIN_RECP_FL_1; 1.
 DR PROSITE: PS0262; G-PROTEIN_RECP_FL_2; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein;
 KW Phosphorylation; Lipoprotein; Palmitate.
 FT DOMAIN 1 37 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 38 63 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 64 72 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 73 93 2 (POTENTIAL).
 FT DOMAIN 94 118 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 119 140 3 (POTENTIAL).
 FT DOMAIN 141 163 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 164 183 4 (POTENTIAL).
 FT DOMAIN 184 191 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 192 211 5 (POTENTIAL).
 FT DOMAIN 212 240 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 241 266 6 (POTENTIAL).
 FT DOMAIN 267 279 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 280 300 7 (POTENTIAL).
 FT DOMAIN 301 317 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 29 29 N-LINKED (GLCNAc...) (POTENTIAL).
 FT LIPID 315 315 PALMITATE (POTENTIAL).
 SQ SEQUENCE 317 AA; 34699 MW; 6615D214E1D247F CRC64;

Query Match 2.8%; Score 8; DB 1; Length 317;
 Best Local Similarity 100.0%; Pred. No. 6.7;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 137 RARRAIA 144
 |||||
 DB 160 RARRAIA 167

RESULT 9
 FDBH_METUA STANDARD; PRT; 379 AA.
 ID FDBH_METUA
 AC Q60316;

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Putative formate dehydrogenase beta chain (EC 1.2.1.2).
GN M00005.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
OC Methanocaldococcaceae; Methanocaldococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Sult C.J., White O., Olsen G.J., Zhou L., Fleischman R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhman J.L., Nguyen D.,
RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
RA Klenk H.-P., Frazer C.M., Smith H.O., Woese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
RT jannaschii";
RL Science 273:1058-1073(1996).
CC -1- CATALYTIC ACTIVITY: Formate + NAD(+) -> CO(2) + NADH.
CC -1- COFACTOR: BINDS TWO 4FE-4S CLUSTERS (PROBABLE).
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS (BY SIMILARITY).
CC -1- SIMILARITY: THE IRON-SULFUR CENTERS ARE SIMILAR TO THOSE OF
CC 'BACTERIAL-TYPE' 4FE-4S FERREDOXINS.
CC -1- SIMILARITY: STRONG, TO THE BETA SUBUNIT OF M.THERMOAUTOTROPHICUM
CC FDH.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U67459; AAB97986.1; -
DR TIGR: M00005; -
DR InterPro: IPR001450; 4Fe4S_ferredoxin.
DR Pfam: PF00037; Fer4; 1.
DR PROSITE: PS00198; 4Fe4S_FERREDOXIN; 2.
KW Hypothetical protein; Oxidoreductase; NAD; Electron transport;
KW Iron-sulfur; 4Fe-4S; Complete proteome.
KM METAL 280
FT METAL 283 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 286 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 290 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 290 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 330 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 333 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 336 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
FT METAL 340 IRON-SULFUR (4FE-4S) (BY SIMILARITY).
SQ SEQUENCE 379 AA; 43014 MM; 9C257CCAD5547F5A CRC64;

Query Match 2.8%; Score 8; DB 1; Length 379;
Best Local Similarity 100.0%; Pred. No. 7.8;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

DE Zona pellucida sperm-binding protein 3 precursor (Zona pellucida
DE glycoprotein zp3) (Sperm receptor) (Zona pellucida protein C).
GN zp3.
OS Mesocricetus auratus (Golden hamster).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Mesocricetus.
OX NCBI_TaxID=10036;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=ovary;
RX MEDLINE=91078540; PubMed=2257975;
RA Kinloch R.A., Ruiz-Seller B., Wasserman P.M.;
RT "Genomic organization and polypeptide primary structure of zona
RT pellucida glycoprotein zp3, the hamster sperm receptor.";
RL Dev. Biol. 142:414-421(1990).
CC -1- FUNCTION: FUNCTIONS AS A SPERM-RECEPTOR. IT IS RESPONSIBLE FOR
CC SPERM-ADHESION TO THE ZONA PELLUCIDA, AND MAY CONTRIBUTE TO THE
CC SPECIES-SPECIFICITY OF THE INSEMINATION.
CC -1- SUBUNIT: ZP3 FORMS WITH ZP1 AND ZP2 THE ZONA PELLUCIDA, IN
CC WHICH ZP2 AND ZP3 COMPLEX INTO COPOLYMERS CROSS-LINKED BY ZP1.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Extracellular
CC matrix.
CC -----
CC -1- TISSUE SPECIFICITY: OOCYTES.
CC -1- DEVELOPMENTAL STAGE: GROWING OOCYTES.
CC -1- PTM: SULFATED GLYCOPROTEIN WITH O-LINKED OLIGOSACCHARIDES.
CC -1- SIMILARITY: CONTAINS 1 ZP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M63629; AAA37079.1; -
DR InterPro: IPR001507; Endoglin/CD105.
DR Pfam: PF00100; zona_pellucida; 1.
DR PRINTS: PR00023; ZPELLUCIDA.
DR SMART: SM00241; zp; 1.
DR PROSITE: PS00682; ZP_DOMAIN; 1.
KW Glycoprotein; Signal; Sulfation; Sperm; Receptor; Transmembrane;
KW Extracellular matrix.
KM SIGNAL 1
FT CHAIN 23 422
FT DOMAIN 23 386 EXTRACELLULAR (POTENTIAL).
FT TRANSMEM 387 407 POTENTIAL.
FT DOMAIN 408 422 CYTOPLASMIC (POTENTIAL).
FT DOMAIN 45 306 PRO-RICH.
FT DOMAIN 119 158 N-LINKED (GLCNAC...) (POTENTIAL).
FT DOMAIN 208 257 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 146 146 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 271 271 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD 302 302 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 422 AA; 45827 MM; D0F95BE7F8B7E01 CRC64;

Query Match 2.8%; Score 8; DB 1; Length 422;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

DE Transport ATP-binding protein cydc.
GN CYDC OR H1156.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Kesteven K., Sutton G., Fitch W., Fields C.A., Merrick J.M.,
RA McCreary K., Sutton G., Fitch W., Fields C.A., Merrick J.M.,
RA Scott J.D., Shiley R., Liu L.-I., Glodok A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uterback T.R., Hanna M.C., Nguyen D.T., Sauder D.M., Brandon R.C.,
RA Fine L.D., Fitchman J.L., Fuhrman J.L., Geoghegan N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae
RD Science 269:496-512(1995).
CC -I- FUNCTION: SOMEHOW INVOLVED IN THE CYTOCHROME D BRANCH OF AEROBIC
CC RESPIRATION. SEEMS TO BE A COMPONENT OF A TRANSPORT SYSTEM
CC (BY SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Probable).
CC -I- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. MSBA SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U32795; AAC22811.1; -.
DR TIGR: H1156; -.
DR InterPro: IPR003593; AAA_ATPase.
DR InterPro: IPR003439; ABC_transportr.
DR InterPro: IPR001140; ABCtransportr.
DR Pfam: PF00005; ABC_tran; 1.
DR Pfam: PF00664; ABC_membrane; 1.
DR PRODOM: PD000006; ABC_transportr; 1.
DR SMART: SM00382; AAA; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER; 1.
KW ATP-binding; Transport; Transmembrane; Inner membrane;
KW Complete proteome.
FT TRANSMEM 16 36 POTENTIAL.
FT TRANSMEM 38 58 POTENTIAL.
FT TRANSMEM 133 153 POTENTIAL.
FT TRANSMEM 155 175 POTENTIAL.
FT TRANSMEM 244 264 POTENTIAL.
FT TRANSMEM 281 301 POTENTIAL.
FT NP_BIND 372 379 ATP (POTENTIAL).
SQ SEQUENCE 576 AA; 64831 MW; A9ACDB9B29441B3 CRC64;

```

```

Query Match 2.8%; Score 8; DB 1; Length 576;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 12
GGT5_HUMAN STANDARD; PRT; 586 AA.
AC P36269; 096PCL;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)

```

```

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Gamma-glutamyltranspeptidase 5 precursor (EC 2.3.2.2) (Gamma-
GN glutamyltransferase 5) (GGT-Rel).
GN GGT1A1 OR GGT5.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Placenta;
RX MEDLINE=91296809; PubMed=1676842;
RA Heisterkamp N., Rajpert-De Meyts E., Uribe L., Forman H.J.,
RA Groffen J.;
RT "Identification of a human gamma-glutamyl cleaving enzyme related to,
RT but distinct from, gamma-glutamyl transpeptidase."
RT Proc. Natl. Acad. Sci. U.S.A. 88:6303-6307(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RX Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
CC -I- FUNCTION: CAN HYDROLYZE THE GAMMA-GLUTAMYL MOIETY OF GLUTATHIONE;
CC AS WELL AS CONVERT LEUKOTRIENE C4 TO LEUKOTRIENE D4.
CC -I- CATALYTIC ACTIVITY: (5-L-glutamyl)-peptide + an amino acid -
CC peptide + 5-L-glutamyl-amino acid.
CC -I- SUBUNIT: THIS ENZYME CONSISTS OF TWO POLYPEPTIDE CHAINS, WHICH ARE
CC SYNTHESIZED IN PRECURSOR FORM FROM A SINGLE POLYPEPTIDE (BY
CC SIMILARITY).
CC -I- SUBCELLULAR LOCATION: Type II membrane protein (Probable).
CC -I- SIMILARITY: STRONG, TO OTHER GGT ENZYMES AND TO GL-7ACA ACYLASES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M64099; AA58503.1; -.
DR EMBL: BC011362; AAH11362.1; -.
DR PIR: A41125; A41125.
DR MEROPS: T03.002; -.
DR GeneW: HGNC:4260; GGT1A1.
DR MIM: 137168; -.
DR InterPro: IPR00101; Gglutnsptdase.
DR Pfam: PF01019; G-glu-transpept; 1.
DR PRINTS: PR01210; GGTTRANSPASE.
DR PROSITE: PS00462; G-GLU-TRANSPPEPTIDASE; 1.
KW Transferrase; Acyltransferase; Zymogen; Glycoprotein; Transmembrane;
KW Glutathione biosynthesis; Signal-anchor.
FT CHAIN 1 387 GAMMA-GLUTAMYLTRANSPPEPTIDASE 5 HEAVY
FT CHAIN 388 586 GAMMA-GLUTAMYLTRANSPPEPTIDASE 5 LIGHT
FT FT CHAIN (POTENTIAL).
FT FT CHAIN (POTENTIAL).
FT FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT TRANSMEM 7 29 (PROBABLE).
FT DONAIN 30 586 LOMENAL (POTENTIAL).
FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 204 204 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 347 347 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 535 535 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 550 550 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 330 330 R -> K (IN REF. 2).
FT CONFLICT 437 437 W -> R (IN REF. 2).
SQ SEQUENCE 586 AA; 62319 MW; 1B5543CB0934B16B CRC64;

```

```

Query Match 2.8%; Score 8; DB 1; Length 586;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 61 LGIGLALA 68
 DB 14 LGIGLALA 21

RESULT 13

HSP1_SAGIM STANDARD; PRT; 49 AA.

AC P24714;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sperm protamine P1 (Cysteine-rich protamine).
 RN PRM1.

OS Saguinus imperator (Tamarin).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae; Saguinus.
 OX NCBI_TaxID=9491;

RN
 RP SEQUENCE FROM N.A.

RC TISSUE=Liver;
 RX MEDLINE=92051332; PubMed=1840669;

RA Queralt R., Oliva R.;
 RT "Protamine 1 gene sequence from the primate Saguinus imperator isolated with PCR using consensus oligonucleotides.";

RL Nucleic Acids Res. 19:5786-5786(1991).

CC -1- FUNCTION: PROTAMINES SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.

CC -1- SUBUNIT: CROSS-LINKED BY INTERCHAIN DISULFIDE BONDS AROUND THE DNA-HELIX (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Nuclear.

CC -1- TISSUE SPECIFICITY: TESTIS.

CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P1 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL: X61678; CAA43853.1; -

DR PIR: S22582; S22582.

DR InterPro: IPR000221; Protamine_P1.

DR Pfam: PF00260; protamine_P1.1.

DR PROSITE: PS00048; PROTAMINE_P1.1.

KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding; Tests; DNA condensation; Nuclear protein.

KW Tests; DNA condensation; Nuclear protein.

FT INIT_MET 0
 FT SEQUENCE 49 AA; 6545 MW; 8389C403F5B207F6 CRC64;

OY 42 ORRRGR 48
 DB 17 ORRRGR 23

RESULT 14

HSP1_DIDMA STANDARD; PRT; 57 AA.

AC P35305;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Sperm protamine P1.

DE Sperm protamine P1.

RN PRM1.

OS Didelphis marsupialis virginiana (North American opossum), and Monodelphis domestica (Short-tailed grey opossum).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Metatheria; Didelphimorphia; Didelphidae; Didelphis.
 OX NCBI_TaxID=9267, 13616;

RESULT 15

A62F_DROME STANDARD; PRT; 115 AA.

AC 046202;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Accessory gland protein Acp62F precursor.
 DE ACP62F OR CG1262.

OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;

RN [1]

RP SEQUENCE FROM N.A., FUNCTION, AND TISSUE SPECIFICITY.

RC STRAIN=Canton-S; TISSUE=Male accessory gland;
 RX MEDLINE=98135120; PubMed=9474779;

CC -1- TISSUE SPECIFICITY: TESTIS.

CC -1- SIMILARITY: BELONGS TO THE PROTAMINE P1 FAMILY.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL: U35448; CAA52193.1; -

DR EMBL: U35448; CAA52193.1; -

DR PIR: S34045; S34045.

DR InterPro: IPR000221; Protamine_P1.

DR Pfam: PF00260; protamine_P1.1.

DR PROSITE: PS00048; PROTAMINE_P1.1.

KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding; Tests; DNA condensation; Nuclear protein.

KW Tests; DNA condensation; Nuclear protein.

FT INIT_MET 0
 FT SEQUENCE 57 AA; 7810 MW; 283715B280214E52 CRC64;

OY 43 RRRGRG 49
 DB 34 RRRGRG 40

Query Match 2.5%; Score 7; DB 1; Length 57;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RA Wolfner M.F., Harada H.A., Bertram M.J., Stelick T.J., Kraus K.W.,
 RA Kalb J.M., Lung Y.O., Neubaum D.M., Park M., Trim U.K.;
 RT "New genes for male accessory gland proteins in *Drosophila*
 RT melanogaster.";
 RL Insect Biochem. Mol. Biol. 27:825-834(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champ M., Pfeiffer B.D.,
 RA Wen K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.C., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Rainett K., Remington K., Saunders R.D.C., Scheefel F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 RN [3]
 RP SEQUENCE OF 7-111 FROM N.A.
 RC STRAIN-Various strains;
 RX MEDLINE=20556153; PubMed=11102381;
 RA Begun D.J., Whitley P., Todd B.L., Waldrup-Dall H.M., Clark A.G.;
 RT "Molecular population genetics of male accessory gland proteins in
 RT *Drosophila*.";
 RL Genetics 156:1879-1888(2000).
 CC -1- FUNCTION: RESPONSIBLE FOR PHYSIOLOGICAL AND BEHAVIORAL CHANGES IN
 CC MATED FEMALE FLIES. MAY CONTRIBUTE TO THE TOXICITY OF SEMINAL
 CC FLUID AND THE DECREASED LIFE-SPAN OF MATED FEMALES. MAY ALSO
 CC AFFECT NEUROMUSCULAR EVENTS AFTER MATING CONCERNING SPERM STORAGE
 CC AND EGG RELEASE.
 CC -1- SUBCELLULAR LOCATION: Secreted (Probable).
 CC -1- TISSUE SPECIFICITY: SEMINAL FLUID.
 CC -1- SIMILARITY: SOME, TO P. NIGRIVENTER TX2-6.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL, U085763; AAB96387.1; -.

DR EMBL: AE003475; AAF47683.1; -
 DR EMBL: AY010608; AAG35367.1; -
 DR EMBL: AY010609; AAG35368.1; -
 DR EMBL: AY010610; AAG35369.1; -
 DR EMBL: AY010611; AAG35370.1; -
 DR EMBL: AY010612; AAG35371.1; -
 DR EMBL: AY010613; AAG35372.1; -
 DR EMBL: AY010614; AAG35373.1; -
 DR EMBL: AY010615; AAG35374.1; -
 DR EMBL: AY010616; AAG35375.1; -
 DR EMBL: AY010617; AAG35376.1; -
 DR FLYBase: FBgn0020509; ACP62F.
 DR InterPro: IPR002919; T1L_Cysrich.
 DR Pfam: PF01826; T1L; 1.
 KW Behavior; signal.
 FT SIGNAL 1 24
 FT CHAIN 25 115
 FT SEQUENCE 115 AA; 12570 MW; 4326A6F6C32291D CRC64;
 SQ
 Query Match 2.5%; Score 7; DB 1; Length 115;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 68 ACIGLL 74
 Db 11 ACIGLL 17
 |||||
 RESULT 16
 Y151_ECOLI
 ID Y151_ECOLI STANDARD; PRT; 118 AA.
 AC P03838;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Insertion element IS5 very hypothetical 12 kda protein.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia.
 OX NCBI_Taxid=562;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82028653; PubMed=6269959;
 RA Engler J.A., van Bree M.P.;
 RT "The nucleotide sequence of IS5 from *Escherichia coli*.";
 RL Gene 14:165-174(1981).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=82028652; PubMed=6269958;
 RA Engler J.A., van Bree M.P.;
 RT "The nucleotide sequence and protein-coding capability of the
 RT transposable element IS5.";
 RL Gene 14:155-163(1981).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
 RA Davis K., Federspiel N., Hyman R., Kalman S., Komp C., Kurdi O.,
 RA Ishikari D., Lew H., Lin D., Namath A., Oetner P., Roberts D.,
 RA Davis R.W.;
 RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
 RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R.,
 RA Ichihara S., Mizuno T., Makino K., Nakata A., Yura T., Sempel G.,
 RA Mizobuchi K.;
 RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN-K12;
 RX MEDLINE=97251357; PubMed=9097039;
 RA Alba H., Bada T., Fujita K., Hayashi K., Inada T., Isono K.,

```

RA Itoh T., Kasai H., Kashimoto K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizouchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nishimoto H., Nishio Y.,
RA Oshima T., Saito N., Sampel G., Seki Y., Sivasubraman S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horieuchi T.:
RT "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28,040,1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J01734: ? NOT_ANNOTATED_CDS.
DR EMBL: U70214: AAB086B0.1: -.
DR EMBL: D83536: ? NOT_ANNOTATED_CDS.
DR EMBL: D90771: BAA14925.1: -.
DR EMBL: D90772: BAA14935.1: -.
DR EMBL: D90831: BAA15715.1: -.
DR EMBL: D90841: BAA15872.1: -.
DR EMBL: D90847: BAA15958.1: -.
DR EMBL: D90848: BAA15963.1: -.
DR PIR: A04466: IECC5B.
KW Hypothetical protein; Transposable element.
SQ SEQUENCE 118 AA; 12270 MW; 348014FAC765058E CRC64;

Query Match 2.5%; Score 7; DB 1; Length 118;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 23 DCGAVRQ 29
DB 107 DCGAVRQ 113

RESULT 17
ID IL13_MOUSE STANDARD; PRT; 131 AA.
AC P20109;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Interleukin-13 precursor (IL-13) (T-cell activation protein p600).
GN IL13 OR IL-13.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_Taxid=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=89093958; PubMed=2521353;
RA Brown K.D., Zurawski S.M., Mosmann T.R., Zurawski G.:
RT "A family of small inducible proteins secreted by leukocytes are
RT members of a new superfamily that includes leukocyte and
RT fibroblast-derived inflammatory agents, growth factors, and
RT indicators of various activation processes.";
RL J. Immunol. 142:679-687(1989).
CC -1- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION.
CC SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.
CC MAY BE CRITICAL IN REGULATING INFLAMMATORY AND IMMUNE RESPONSES
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M23504: AAA40149.1: -.
DR PIR: E30552: E30552.
DR HSSP: P35225: 3ITR.
DR MGI: MGI:96541: IL13.
DR InterPro: IPR003634; Interleukin_13.
DR InterPro: IPR001325; Interleukin_4_13.
DR Pfam: PF03487; Interleukin_13; 1.
DR ProDom: PD015987; Interleukin_13; 1.
DR SMART: SM00190; IL4_13; 1.
DR PROSITE: PS00838; INTERLEUKIN_4_13; 1.
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 1 131 INTERLEUKIN_13.
FT DISULFID 51 79 BY SIMILARITY.
FT DISULFID 67 93 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 52 52 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 75 75 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 131 AA; 14107 MW; 954F93F105713FED CRC64;

Query Match 2.5%; Score 7; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACLG 71
DB 9 LALACLG 15

RESULT 18
ID IL13_RAT STANDARD; PRT; 131 AA.
AC P42203;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Interleukin-13 precursor (IL-13) (T-cell activation protein p600).
GN IL13 OR IL-13.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Kidney cortex;
RX MEDLINE=94092138; PubMed=7916615;
RA Lakkis F.G., Cruet E.N.:
RT "Cloning of rat Interleukin-13 (IL-13) cDNA and analysis of IL-13
RT gene expression in experimental glomerulonephritis.";
RL Biochem. Biophys. Res. Commun. 197:612-618(1993).
CC -1- FUNCTION: CYTOKINE. INHIBITS INFLAMMATORY CYTOKINE PRODUCTION.
CC SYNERGIZES WITH IL2 IN REGULATING INTERFERON-GAMMA SYNTHESIS.
CC MAY BE CRITICAL IN REGULATING INFLAMMATORY AND IMMUNE RESPONSES
CC (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- SIMILARITY: BELONGS TO THE IL-4 / IL-13 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U26913: AAA16478.1: -.
DR HSSP: P35225: 3ITR.
DR InterPro: IPR003634; Interleukin_13.
DR InterPro: IPR001325; Interleukin_4_13.

```

```

DR Pfam: PF03487; Interleukin_13; 1.
DR Prodom: PD015987; Interleukin_13; 1.
DR SMART: SM00190; IL4_13; 1.
DR PROSITE: PS00838; INTERLEUKIN_4_13; 1.
KW Cytokine; Glycoprotein; Signal.
FT SIGNAL 1 21 BY SIMILARITY.
FT CHAIN 22 131 INTERLEUKIN-13.
FT DISULFID 52 80 BY SIMILARITY.
FT DISULFID 68 94 BY SIMILARITY.
FT CARBOHYD 42 42 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 53 53 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 76 76 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 121 121 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 131 AA; 14093 MW; E508CAB8DE8C201 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 131;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 LALACLG 71
DB 9 LALACLG 15

RESULT 19
ID YK01_PYRHO STANDARD; PRT; 147 AA.
AC 057781;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE Hypothetical protein PH2001.
GN PH2001.
OS Pyrococcus horikoshii.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=53953;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OT3;
RA MEDLINE=98344137; PubMed=9679194;
RA Kawarabayashi Y., Sawada M., Horikawa H., Hino Y.,
RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamaya M., Ohnuki Y.,
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki Y., Kishida N., Oguchi A.,
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
RA Masuchi Y., Shizuza H., Kikuchi H.;
RT "Complete sequence and gene organization of the genome of a hyper-
thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
RL DNA Res. 5:55-76(1998).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP000001; BAA31940.1; -
DR EMBL: AP000007; BAA31940.1; JOINED.
DR EMBL: AP000007; BAA31943.1; -
DR EMBL: AP000001; BAA31943.1; JOINED.
KW Hypothetical protein; Transmembrane; Complete proteome.
FT TRANSMEM 41 61 POTENTIAL.
FT TRANSMEM 67 87 POTENTIAL.
SQ SEQUENCE 147 AA; 15324 MW; 247ED12FCECF265B9 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 LACIGLL 73

```

```

DB 45 LACIGLL 51

RESULT 20
ID TNFC_PIG STANDARD; PRT; 150 AA.
AC O9TSV8;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
DE necrosis factor ligand superfamily member 3) (Fragment).
DE LTB OR TNFSF3 OR TNFC.
OS Sus scrofa (pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=large white; Tissue=Fibroblast;
RX MEDLINE=21108615; PubMed=11169259;
RA Chardon P., Rogel-Galliard C., Cattolico L., Duprat S., Vaiman M.,
RA Renard C.;
RT "Sequence of the swine major histocompatibility complex region
RT containing all non-classical class I genes."
RL Tissue Antigens 57:55-65(2001).
CC -I- FUNCTION: Cytokine that binds to LTB/ TNFSF3. May play a specific
CC role in immune response regulation. Provides the membrane anchor
CC for the attachment of the heterotrimeric complex to the cell
CC surface.
CC -I- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or
CC (less prevalent) two LTA and one LTB subunits.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein (by similarity).
CC -I- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ251914; CAB63851.1; -
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF; 1.
DR PRINTS: PR01234; TNFCROSISFCT.
DR Prodom: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS00049; TNF_2; 1.
KW Cytokine; Cytotoxin; Glycoprotein.
FT NON-TER 1 1
FT CARBOHYD 128 128 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 150 AA; 16423 MW; F5C4C657658B48 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 150;
Best Local Similarity 100.0%; Pred. No. 32;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 GLYLYYC 200
DB 37 GLYLYYC 43

RESULT 21
ID RA05_ORYSA STANDARD; PRT; 157 AA.
AC Q01881;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)

```

```

DT 01-JUN-1994 (rel. 29, last annotation update)
DE Seed allergenic protein RA5 precursor.
GN RA5.
OS Oryza sativa (Rice).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: Liliopsida: Poales: Poaceae:
OC Ehrhartoideae: Oryzaceae: Oryza.
OX NCBI_TaxID=4530;
RN RP SEQUENCE FROM N.A.
RC TISSUE=Seed:
RA MEDLINE=93144699; PubMed=7678765;
RA Adachi T., Izumi H., Yamada T., Tanaka K., Takeuchi S.,
RA Nakamura R., Matsuda T.;
RT "Gene structure and expression of rice seed allergenic proteins
RT belonging to the alpha-amylase/trypsin inhibitor family.";
RL Plant Mol. Biol. 21:239-248(1993).
CC -1- FIVE DISULFIDE BONDS ARE PRESENT (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE CEREAL TRYPSTIN/ALPHA-AMYLASE INHIBITOR
CC FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D11430; BAA01996.1; -
DR PIR: S31078; S31078.
DR HSSP: P01085; IHSS.
DR InterPro: IPR003612; AAI.
DR InterPro: IPR001768; TRY/amy1_inhbr.
DR Pfam: PF00234; tryp_alpha_aml1; 1.
DR PRINTS: PRO0808; AMLASEINHBR.
DR SMART: SM00499; AAI; 1.
DR PROSITE: PS00426; CEREAL_TRYP_AMYL_INH; 1.
KW Allergen; Multigene family; Signal.
FT SIGNAL 1 26 POTENTIAL.
FT CHAIN 27 157 SEED ALLERGENIC PROTEIN RA5.
SQ SEQUENCE 157 AA; 17118 MW; CBA5495FBFB399E6 CRC64;
Query Match 2.5%; Score 7; DB 1; Length 157;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 72 LLLAVVS 78
DB 11 LLLAVVS 17
RESULT 22
HLPA_ECOLI STANDARD: PRT; 161 AA.
AC P11457;
DT 01-OCT-1989 (rel. 12, Created)
DT 01-OCT-1989 (rel. 12, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE Histone-like protein HLP-1 precursor (DNA-binding 17 kDa protein).
GN HLPA OR SKP OR OMPH OR B0178 OR 20190 OR ECS0180.
OS Escherichia coli, and
OS Escherichia coli O157:H7.
OC Bacteria: Proteobacteria; gamma subdivision; Enterobacteriaceae:
OC Escherichia.
OX NCBI_TaxID=562, 83334;
RN RP SEQUENCE FROM N.A., AND SEQUENCE OF 21-40.
RX MEDLINE=88329735; PubMed=2843433;
RA Holck A., Kleppe K.;
RT "Cloning and sequencing of the gene for the DNA-binding 17k protein
RT of Escherichia coli.";
RL Gene 67:117-124(1988).

```

```

RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RL "The complete genome sequence of Escherichia coli K-12.";
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / W3110;
RA Takemoto K., Mori H., Murayama N., Kataoka K., Yano M., Itoh T.,
RA Yamamoto Y., Inokuchi H., Miki T., Hatada E., Fukuda R., Ichihara S.,
RA Mizuno T., Makino K., Nakata A., Yura T., Sempel G., Mizobuchi K.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of the
RT 4.0 - 6.0 min (189,987 - 281,416bp) region.";
RL Submitted (FEB-1996) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Schramm S., Duncan M., Allen E., Araujo R., Aparicio A., Chung E.,
RA Davis K., Federspiel N., Hyman R., Kaiman S., Komp C., Kurdi O.,
RA Lashari D., Lew H., Lin D., Namath A., Oefner P., Roberts D.,
RA Davis R.W.;
RL Submitted (SEP-1996) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RX MEDLINE=21074935; PubMed=11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Postel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,
RA Grobbeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamousis K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RT "Genome sequence of enterohemorrhagic Escherichia coli O157:H7.";
RL Nature 409:529-533(2001).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RX MEDLINE=21156231; PubMed=11258796;
RA Hayashi T., Makino K., Onishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka H., Tooe T.,
RA Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T.,
RA Kuhara S., Shiba T., Hattori M., Shingawa H.;
RT "Complete genome sequence of enterohemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
RN [7]
RP SEQUENCE OF 72-161 FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=91100302; PubMed=1987124;
RA Dicker I.B., Seetharam S.R.;
RT "Cloning and nucleotide sequence of the fira gene and the fira200(79)
RT allele from Escherichia coli.";
RL J. Bacteriol. 173:334-344(1991).
RN [8]
RP SEQUENCE OF 21-32.
RC STRAIN-K12 / EMG2;
RX MEDLINE=97443975; PubMed=9298646;
RA Link A.J., Robison K., Church G.M.;
RT "Comparing the predicted and observed properties of proteins encoded
RT in the genome of Escherichia coli K-12.";
RL Electrophoresis 18:1259-1313(1997).
RN [9]
RP SIMILARITY TO S.TYPHIMURIUM OMPH.
RX MEDLINE=90201355; PubMed=2318304;
RA Hirvas L., Coleman J., Koski P., Vaara M.;
RT "Bacterial 'histone-like protein I' (HLP-I) is an outer membrane
RT constituent?";
RL FEBS Lett. 262:123-126(1990).
CC -1- SUBUNIT: HOMOTETRAMER.

```

CC -1- SUBCELLULAR LOCATION: EITHER IN THE NUCLEOID (CHROMATIN) OR IN
CC THE OUTER MEMBRANE.
CC -1- SIMILARITY: BELONGS TO THE OMPH/HMPA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: M21118; AAA24630.1; -
DR EMBL: AE000127; AAC73289.1; -
DR EMBL: D83536; BAA77853.1; -
DR EMBL: U70214; AAB08607.1; -
DR EMBL: AE005194; AAG54460.1; -
DR EMBL: AP002550; BAB33603.1; -
DR EMBL: X54797; CAA38567.1; -
DR EMBL: X75465; CAA53207.1; -
DR PIR: JTO304; DNEC17.
DR PIR: S13728; S13728.
DR Ecogene; BG10455; hnpa.
DR DNA-binding; Outer membrane; Signal; Complete proteome.
KW SIGNAL 1 20
FT CHAIN 21 161 HISTONE-LIKE PROTEIN HLP-1.
SQ SEQUENCE 161 AA; 17688 MW; 2A966BB083FE675 CMC64;

Query Match 2.5%; Score 7; DB 1; Length 161;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 62 GLGLALA 68
Db 9 GLGLALA 15

RESULT 23
YLMA_CAEEL
ID YLMA_CAEEL STANDARD; PRT; 170 AA.
AC P34378;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein D2007.4 in chromosome III.
GN D2007.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilita; Rhabditoidea;
OC Rhabdilitae; Pezodermiae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol N2;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Johnston L., Jones M., Kershaw J., Kirsten J., Lalister N.,
RA Latreille P., Lightning J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Showken R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sulston J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterston R., Watson A., Weinstock L., Wilkinson-Spratt J.,
RA Womlidan P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
elegans";
RL Nature 368:32-38(1994).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: L16560; AAA27999.1; -
DR PIR: S44789; S44789.
DR Wormpep; D2007.4; CE00129.
KW Hypothetical protein.
SQ SEQUENCE 170 AA; 19396 MW; 22301D7C65638135 CMC64;

Query Match 2.5%; Score 7; DB 1; Length 170;
Best Local Similarity 100.0%; Pred. No. 36;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 221 VLALRCL 227
Db 94 VLALRCL 100

RESULT 24
LACB_BUBBU
ID LACB_BUBBU STANDARD; PRT; 180 AA.
AC P02735; O62822;
DT 21-JUL-1986 (Rel. 01, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-lactoglobulin precursor (Beta-LG).
GN LGB.
OS Bubalus bubalis (Domestic water buffalo).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bubalus.
OX NCBI_TaxID=89462;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Mammary gland;
RX MEDLINE=99304500; PubMed=10376212;
RA Das P., Jain S., Nayak S., Apparo K.B.C., Torey S.M., Garg L.C.;
RT "Molecular cloning and sequence analysis of the cDNA encoding
RT beta-lactoglobulin in Bubalus bubalis";
RL DNA Seq. 10:105-108(1999).
RN [2]
RP SEQUENCE OF 19-180.
RA Kolde H.-J., Libertori J., Brauntzer G.;
RT "The amino acid sequence of the water buffalo beta-lactoglobulin";
RL Milchwissenschaft 36:83-86(1981).
CC -1- FUNCTION: Primary component of whey, it binds retinol and is
CC probably involved in the transport of that molecule.
CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS
CC AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Synthesized in mammary gland and secreted
CC in milk.
CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.
CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

DR EMBL: AJ005429; CAA06532.1; -
DR PIR: A03219; IGBUI.
DR HSSP; P02754; IBSQ.
DR InterPro; IPR002345; Lipocalin.
DR InterPro; IPR000566; Lipocalin_cytrABP.
DR Pfam; PF00061; Lipocalin_1.
DR PRINTS; PR00179; LIPOCALIN.
DR PROSITE; PS00213; LIPOCALIN; 1.
KW Milk; Whey; Retinol-binding; Transport; Lipocalin; Signal.


```

FT SIGNAL 1 18
FT CHAIN 19 180
FT DISULFID 84 178
FT DISULFID 124 137
FT DISULFID 124 139
SQ SEQUENCE 180 AA; 20023 MW; 6836C97B2C2E33CF CRC64;
ALTERNATE:
Query Match 2.5%; Score 7; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGALAC 69
DB 8 LGALAC 14

RESULT 25
LACB_CAPHI STANDARD; PRT: 180 AA.
ID LACB_CAPHI
AC P02756;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Beta-lactoglobulin precursor (beta-Lg).
GN LGB.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ssp. aegagrus; TISSUE=Mammary gland;
RX MEDLINE=94042559; PubMed=8226387;
RA Folch J.M., Coll A., Sanchez A.;
RT "Cloning and sequencing of the cDNA encoding goat beta-
RT lactoglobulin."
RL J. Anim. Sci. 71:2832-2832(1993).
RN [2]
RP SEQUENCE FROM N.A.
RA Kim J., Kim A., Kim J., Yu M.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=95213451; PubMed=7699130;
RA Folch J., Coll A., Sanchez A.;
RT "Complete sequence of the caprine beta-lactoglobulin gene."
RL J. Dairy Sci. 77:3493-3497(1994).
RN [4]
RP SEQUENCE OF 19-180.
RX MEDLINE=80070611; PubMed=511095;
RA Preaux G., Brauntz G., Schrank B., Stangl A.;
RT "The amino acid sequence of goat beta-lactoglobulin."
RL Hoppe-Seyler's Z. Physiol. Chem. 360:1595-1604(1979).
CC -1- FUNCTION: Primary component of whey, it binds retinol and is
CC probably involved in the transport of that molecule.
CC -1- SUBUNIT: UNDER PHYSIOLOGICAL CONDITIONS BETA-LACTOGLOBULIN EXISTS
CC AS AN EQUILIBRIUM MIXTURE OF MONOMERIC AND DIMERIC FORMS.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Synthesized in mammary gland and secreted
CC in milk.
CC -1- MISCELLANEOUS: ALTERNATE DISULFIDE BONDS OCCUR IN EQUAL AMOUNTS.
CC -1- SIMILARITY: BELONGS TO THE LIPOCALIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: X58471; CAA41385.1; -
DR EMBL: Z19569; CAA79623.1; -
DR EMBL: Z19570; CAA79623.1; -
DR EMBL: Z33881; CAA83946.1; -
DR PIR: A03220; IGGT
DR PIR: S14507; S14507.
DR PIR: S14507; S14507.
DR HSSP: P02754; 1B00.
DR InterPro: IPR002345; Lipocalin.
DR InterPro: IPR000566; Lipocalin_cytFABP.
DR pfam: pf00061; Lipocalin_1.
DR PRINTS: PR00179; LIPOCALIN.
DR PROSITE: PS00213; LIPOCALIN; 1.
KW MILK; Whey; Retinol-binding; Transport; Lipocalin; Signal.
FT SIGNAL 1 18
FT CHAIN 19 180
FT DISULFID 84 178
FT DISULFID 124 137
FT DISULFID 124 139
SQ SEQUENCE 180 AA; 19975 MW; C2449BB02A1A80F1 CRC64;
ALTERNATE:
Query Match 2.5%; Score 7; DB 1; Length 180;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 63 LGALAC 69
DB 8 LGALAC 14

RESULT 26
LACB_SHEEP STANDARD; PRT: 180 AA.
ID LACB_SHEEP
AC P02757;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-AUG-1988 (Rel. 08, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE beta-lactoglobulin 1/B, 2/A, and 3/C precursor.
GN beta-lactoglobulin 1/B, 2/A, and 3/C precursor.
OS Ovis aries (Sheep), and
OS Ovis orientalis musimon (Mouflon).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940, 9938;
RN [1]
RP SEQUENCE FROM N.A. (BLG 1 AND 2).
RC SPECIES=Sheep;
RX MEDLINE=88172489; PubMed=3351935;
RA Ali S., Clark A.J.;
RT "Characterization of the gene encoding ovine beta-lactoglobulin.
RT similarity to the genes for retinol binding protein and other
RT secretory proteins."
RL J. Mol. Biol. 199:415-426(1988).
RN [2]
RP SEQUENCE FROM N.A. (BLG 1).
RC SPECIES=Sheep;
RX MEDLINE=87049827; PubMed=3096387;
RA Gave P., Hue-Delahaie D., Mercier J.-C., Soulier S., Vilotte J.-L.,
RA Furet J.-P.;
RT "Ovine beta-lactoglobulin messenger RNA: nucleotide sequence and mRNA
RT levels during functional differentiation of the mammary gland."
RL Biochimie 68:1097-1107(1986).
RN [3]
RP SEQUENCE FROM N.A. (BLG 1).
RC SPECIES=Sheep;
RX MEDLINE=89057492; PubMed=3194215;
RA Harris S., Ali S., Anderson S., Archibald A.L., Clark A.J.;
RT "Complete nucleotide sequence of the genomic ovine beta-lactoglobulin
RT gene."
RL Nucleic Acids Res. 16:10379-10380(1988).
RN [4]
RP SEQUENCE FROM N.A. (BLG 1 AND 2).
RC SPECIES=Sheep;

```



```

CC -----
DR EMBL: AE008763; AAL20376.1; -.
DR EMBL: AL627271; CAD01913.1; -.
DR StycGene: SGT27272; rnfE.
DR InterPro: IPR003667; Rnf_Nqr.
DR Pfam: PF02508; Rnf-Nqr; I.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 34 56 POTENTIAL.
FT TRANSMEM 69 87 POTENTIAL.
FT TRANSMEM 91 113 POTENTIAL.
FT TRANSMEM 126 148 POTENTIAL.
FT TRANSMEM 183 205 POTENTIAL.
SQ SEQUENCE 230 AA: 24318 MW: 51988kDaEAL3F249E CRC64;

Query Match 2.5%; Score 7; DB 1; Length 230;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLA 66
DB 38 ALGLGLA 44

RESULT 28
RNF_ECO57 STANDARD: PRT: 231 AA.
ID RNF_ECO57
AC P58344;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNF OR Z2642 OR ECS2341.
OS Escherichia coli O157:H7.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=83334;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / EDL933 / ATCC 700927;
RA MEDLINE-21074935; PubMed-11206551;
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamowski K.,
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,
RA Welch R.A., Blattner F.R.;
RA "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";
RA Nature 409:529-533(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-O157:H7 / RIMD 0509952;
RA MEDLINE-21156231; PubMed-11258796;
RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tohe T.,
RA Ikeda T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,
RA Kihara S., Shiba T., Hattori M., Shingawa H.;
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli
RT O157:H7 and genomic comparison with a laboratory strain K-12.";
RL DNA Res. 8:11-22(2001).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/

```

```

CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AE005386; AAG56621.1; -.
DR EMBL: AP002558; BAB35764.1; -.
DR InterPro: IPR003667; Rnf_Nqr.
DR Pfam: PF02508; Rnf-Nqr; I.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 1 38 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 39 59 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 60 62 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 63 83 POTENTIAL.
FT TRANSMEM 84 85 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 86 106 POTENTIAL.
FT TRANSMEM 107 124 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 125 145 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 146 181 PERIPLASMIC (POTENTIAL).
FT TRANSMEM 182 202 POTENTIAL.
FT TRANSMEM 203 231 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 231 AA: 24489 MW: 54262kDaD292504C3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLA 66
DB 38 ALGLGLA 44

RESULT 29
RNF_ECOLI STANDARD: PRT: 231 AA.
ID RNF_ECOLI
AC P77179;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNF OR B1632.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RA MEDLINE-9742617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RA "The complete genome sequence of Escherichia coli K-12";
RA Science 277:1453-1474(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / 5751357; PubMed-9097039;
RA MEDLINE-97251357; PubMed-9097039;
RA Alba H., Baba T., Fujita K., Hayashi K., Inada T., Isono K.,
RA Itoh T., Kasa T., Kashiwagi K., Kimura S., Kitakawa M.,
RA Kitagawa M., Makino K., Miki T., Mizobuchi K., Mori H., Mori T.,
RA Motomura K., Nakade S., Nakamura Y., Nasimoto H., Nishio Y.,
RA Oshima T., Saito N., Sempel G., Seki Y., Sivasubramanian S.,
RA Tagami H., Takeda J., Takemoto K., Takeuchi Y., Wada C.,
RA Yamamoto Y., Horikuchi T.;
RA "A 570-kb DNA sequence of the Escherichia coli K-12 genome
RT corresponding to the 28.0-40.1 min region on the linkage map.";
RL DNA Res. 3:363-377(1996).
RN [3]
RP TOPOLOGY.
RC STRAIN-K12 / JM109;
RA MEDLINE-99342054; PubMed-10411911;
RA Saef A., Johansson M., Wallin E., von Heijne G.;
RA "Divergent evolution of membrane protein topology: the Escherichia
RT coli RnfA and RnfE homologues.";

```

```

RL Proc. Natl. Acad. Sci. U.S.A. 96:8540-8544(1999).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AE000258; AAC74704.1; -.
DR EMBL; D90806; BAA15386.1; -.
DR EMBL; D90807; BAA15393.1; -.
DR EMBL; D90808; BAA15416.1; -.
DR Ecogene; EG13938; rnfE.
DR InterPro; IPR003667; Rnf_Nqr.
DR Pfam; PF02508; Rnf_Nqr; 1.
DR Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT DOMAIN 1 38 PERIPLASMIC (POTENTIAL).
FT TRANSSEM 39 59 PERIPLASMIC (POTENTIAL).
FT DOMAIN 60 62 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 63 83 POTENTIAL.
FT DOMAIN 84 85 PERIPLASMIC (POTENTIAL).
FT TRANSSEM 86 106 PERIPLASMIC (POTENTIAL).
FT DOMAIN 107 124 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 125 145 POTENTIAL.
FT DOMAIN 146 181 PERIPLASMIC (POTENTIAL).
FT TRANSSEM 182 202 POTENTIAL.
FT DOMAIN 203 231 CYTOPLASMIC (POTENTIAL).
SQ SEQUENCE 231 AA; 24459 MW; CFA37A2D292604C3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 231;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 ALGLGLA 66
| | | | |
DB 38 ALGLGLA 44

RESULT 30
RNC_COXBU STANDARD; PRT; 233 AA.
ID RNC_COXBU
AC P51837;
DR 01-OCT-1996 (Rel. 34, Created)
DR 01-OCT-1996 (Rel. 34, Last sequence update)
DR 16-OCT-2001 (Rel. 40, Last annotation update)
DE Ribonuclease III (EC 3.1.26.3) (RNase III).
GN RNC.
OS Coxiella burnetii.
OC Bacteria; Proteobacteria; gamma subdivision; Legionellaceae group;
CC Coxiella group; Coxiella.
OX NCBL_TaxID=777;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CB9M1C7;
RX MEDLINE=95131751; PubMed=7830573;
RA Zuber M., Hoover T.A., Powell B.S., Court D.L.;
RT "Analysis of the rnc locus of Coxiella burnetii.";
RL Mol. Microbiol. 14:291-300(1994).
CC -1- FUNCTION: DIGESTS DOUBLE-STRANDED RNA. INVOLVED IN THE PROCESSING
CC OF RIBOSOMAL RNA PRECURSORS AND OF SOME MRNAs (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphomonoester.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: CONTAINS 1 DRBM (DOUBLE-STRANDED RNA-BINDING) DOMAIN.

```

```

CC -1- SIMILARITY: CONTAINS 1 RNASE III DOMAIN.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; L27436; AAA69690.1; -.
DR InterPro; IPR001159; DS_RBD.
DR InterPro; IPR000999; RNase_3.
DR Pfam; PF00035; dsrm; 1.
DR Pfam; PF00636; Ribonuclease_3; 1.
DR SMART; SM00358; DSRM; 1.
DR SMART; SM00353; RIBOC; 1.
DR PROSITE; PS50137; DS_RBD; 1.
DR PROSITE; PS00517; RNASE_3_1; 1.
DR PROSITE; PS50142; RNASE_3_2; 1.
KW Hydrolyase; Nuclease; Endonuclease; RNA-binding.
FT DOMAIN 4 126 RNASE III.
FT DOMAIN 204 220 DRBM.
SQ SEQUENCE 233 AA; 26229 MW; 1A1CB4AD96FD2FF CRC64;

Query Match 2.5%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 10 ARRLPLP 16
| | | | |
DB 164 ARRLPLP 170

RESULT 31
RNF_YERPE STANDARD; PRT; 233 AA.
ID RNF_YERPE
AC Q8ZED4;
DR 15-JUN-2002 (Rel. 41, Created)
DR 15-JUN-2002 (Rel. 41, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNF OR YPO2240.
OS Yersinia pestis.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Yersinia.
OX NCBL_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CO-92 / Biovar Orientalis;
RX MEDLINE=21470413; PubMed=11586360;
RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
RA Prentice M.B., Sebalha M., James K.D., Churcher C., Mungall K.L.,
RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
RA Fellwell T., Hamlin N., Holtroyd S., Jagels K., Karlyshev A.V.,
RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
RA Simmonds M., Skelton J., Stevens K., Whitehead S., Barrett B.G.;
RT "Genome sequence of Yersinia pestis, the causative agent of plague.";
RL Nature 413:523-527(2001).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits: rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AJ414151; CAC91046.1; -

DR InterPro: IPR003667; Rnf.Nqr.

DR Pfam: PF02508; Rnf.Nqr; 1.

KW Electron transport; Transmembrane; Inner membrane; Complete proteome.

FT TRANSMEM 34 56 POTENTIAL.

FT TRANSMEM 69 87 POTENTIAL.

FT TRANSMEM 91 113 POTENTIAL.

FT TRANSMEM 126 148 POTENTIAL.

FT TRANSMEM 184 206 POTENTIAL.

SO SEQUENCE 233 AA; 24587 MW; 491E18F335E8CB90 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 233;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLA 66
| | | | | | | |

DB 38 ALGLGLA 44

RESULT 32

INFE_HAEIN STANDARD; PRT; 235 AA.

ID INFE_HAEIN

AC 057020; P96346;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Electron transport complex protein rnfE.

GN RNF E OR H11688.

OS Haemophilus influenzae.

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;

OC Haemophilus.

OX NCBI_TaxID=727;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-Rd / KW20 / ATCC 51907;

RX MEDLINE=95350630; PubMed=7542800;

RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F., Mckeown A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M., Scott J.D., Shirley R., Liu L.-I., Glodek A., Gocayne J.D., Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D., Usterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C., Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghagen N.S.M., Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O., Venter J.C.;

RA "Whole-genome random sequencing and assembly of Haemophilus influenzae Rnf."

RT Science 269:496-512(1995).

RL -1- FUNCTION: May be part of a membrane complex involved in electron transport (By similarity).

CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC, rnfD, rnfE and rnfG (By similarity).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).

CC -1- SIMILARITY: BELONGS TO THE NORDE/NFAE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: U32841; AAC23334.1; -

DR TIGR: H11688; -

DR InterPro: IPR003667; Rnf.Nqr.

DR Pfam: PF02508; Rnf.Nqr; 1.

DR Electron transport; Transmembrane; Inner membrane; Complete proteome.

FT TRANSMEM 63 83 POTENTIAL.

FT TRANSMEM 93 113 POTENTIAL.

FT TRANSMEM 117 137 POTENTIAL.

FT TRANSMEM 152 172 POTENTIAL.

FT TRANSMEM 206 226 POTENTIAL.

SO SEQUENCE 235 AA; 25845 MW; C054FE96647837A CRC64;

Query Match 2.5%; Score 7; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLGLA 66
| | | | | | | |

DB 62 ALGLGLA 68

RESULT 33

TN14_MOUSE STANDARD; PRT; 239 AA.

ID TN14_MOUSE

AC O9QYH9;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 14.

GN TNFSF14 OR LIGHT.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RX PubMed=10700230;

RA Tamada K., Shinozaki K., Chapoval A.I., Zhu G., Sica G., Flies D., Boone T., Hsu H., Fu Y.-X., Nagata S., Ni J., Chen L.;

RT "Modulation of T-cell-mediated immunity in tumor and graft-versus-host disease models through the LIGHT co-stimulatory pathway."

RL Nat. Med. 6:283-289(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Fetal liver;

RA Misawa K., Nosaka T., Kitamura T., Kojima T.;

RT "Murine LIGHT, a homologue of human LIGHT which is a member of TNF family."

RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Lymphoma;

RA Force W.R., Todd P.K., Mikayama T.;

RT "Mouse LIGHT: molecular genetics, ligand binding and expression."

RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: Cytokine that binds to TNFRSF3/LTR. Binding to the decoy receptor TNFRSF6B modulates its effects. Activates NFkB and stimulates the proliferation of T cells.

CC -1- SUBUNIT: Homotrimer (By similarity).

CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By similarity).

CC -1- PTM: The soluble form derives from the membrane form by proteolytic processing.

CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

DR EMBL: AF123385; AAF76453.1; -

DR AB029155; BAA8559.1; -

DR EMBL: AF227533; AAF36722.1; -

DR HSSP: P01375; 4TSV.

DR MGD: MGI:1355317; Tnfrsf14.

```

DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR PRINTS: PR01234; TNECROSISFCT.
DR ProDom: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS00251; TNF_1; FALSE_NEG.
DR PROSITE: PS00049; TNF_2; 1.
KM Cytokine: Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 239
FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT CHAIN 782 239
FT TUMOR NECROSIS FACTOR FORM.
FT TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY
FT DOMAIN 1 37
FT MEMBER 14, SOLUBLE FORM.
FT TRANSMEM 38 58
FT CYTOPLASMIC (POTENTIAL).
FT SIGNAL-ANCHOR (TYPE II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 59 239
FT EXTRACELLULAR (POTENTIAL).
FT SITE 81 82
FT CLEAVAGE (POTENTIAL).
FT DISULFID 152 187
FT POTENTIAL.
FT CARBOHYD 100 100
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 191 191
FT N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 239 AA; 26338 MW; 217874AC71AD6BE3 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 239;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 121 NLRVPR 127
    |||||
Db 221 NLRVPR 227

RESULT 34
RNF_ESEAE
ID RNF_ESEAE STANDARD; PRT; 240 AA.
AC O9HYB5;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Electron transport complex protein rnfE.
GN RNF_E OR PA3494.
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Hickley M.J., Brinkman F.S.L., Mizoguchi S.D., Warren P.,
RA Garber R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizner J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RA "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen";
RL Nature 406:959-964(2000).
CC -1- FUNCTION: May be part of a membrane complex involved in electron
CC transport (By similarity).
CC -1- SUBUNIT: Composed of at least six subunits; rnfA, rnfB, rnfC,
CC rnfD, rnfE and rnfG (By similarity).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (Potential).
CC -1- SIMILARITY: BELONGS TO THE NORDE/RNFAE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```

CC -----
CC EMBL: AE004770; AAC06882.1;
DR InterPro: IPR003667; Rnf_Ngr.
DR Pfam: PF02508; Rnf_Ngr; 1.
KW Electron transport; Transmembrane; Inner membrane; Complete proteome.
FT TRANSMEM 41 61
FT POTENTIAL.
FT TRANSMEM 71 91
FT POTENTIAL.
FT TRANSMEM 95 115
FT POTENTIAL.
FT TRANSMEM 130 150
FT POTENTIAL.
FT TRANSMEM 184 204
FT POTENTIAL.
SQ SEQUENCE 240 AA; 25794 MW; 3D90687ED462DBB2 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 240;
Best Local Similarity 100.0%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 60 ALGLGLA 66
    |||||
Db 40 ALGLGLA 46

RESULT 35
MOEB_HAEIN
ID MOEB_HAEIN STANDARD; PRT; 243 AA.
AC P45211;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Molybdopterin biosynthesis protein moeb.
GN MOEB OR CHLN OR H11449.
OS Haemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_TaxID=727;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Rd / KW20 / ATCC 51907;
RX MEDLINE=95350630; PubMed=7542800;
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,
RA Karlavague A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,
RA McKenney K., Sulten G., Fitzhugh W., Fields C.A., Gocayne J.D.,
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,
RA Uiterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,
RA Fine L.D., Frichman J.L., Fuhrmann J.L., Geoghagen N.S.M.,
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,
RA Venter J.C.;
RA "Whole-genome random sequencing and assembly of Haemophilus influenzae
RT Rd.";
RL Science 269:496-512(1995).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
CC (MOLYBDOPTERIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE
CC ACTIVATION (MOAD) OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTING
CC FACTOR (MOAF) (BY SIMILARITY).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U32823; AAC23099.1;
DR TIGR: H11449;
DR InterPro: IPR000205; NAD_binding.
DR InterPro: IPR000594; Thif_domain.
DR Pfam: PF00899; Thif; 1.
KW Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 243 AA; 26996 MW; 218A3382A975BBD CRC64;
```

Query Match 2.5%; Score 7; DB 1; Length 243;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNR01 185
 DB 10 LRYNR01 16

RESULT 36
 TNFC_HUMAN STANDARD; PRT: 244 AA.

AC 006643; P78370; Q99761;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
 necrosis factor ligand superfamily member 3).
 GN LTB OR TNFSF3 OR TNFC.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A. (ISOFORM 1), AND PARTIAL SEQUENCE.
 RC TISSUE=T-cell.
 RX MEDLINE=93208881; PubMed=7916655;
 RA Browning J.L., Ngam-Ek A., Lawton P., Demarinis J., Tizard R.,
 Chow E.P., Hession C., O'Brien-Greco B., Foley S.F., Ware C.F.,
 "Lymphotoxin beta, a novel member of the TNF family that forms a
 heteromeric complex with lymphotoxin on the cell surface.";
 RL Cell 72:847-856(1993).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
 RX MEDLINE=97445965; PubMed=9299492;
 RA Watzocha K., Renard N., Charlot C., Bienvu J., Collier B.,
 Salles G.,
 "Identification of two lymphotoxin beta isoforms expressed in human
 lymphoid cell lines and non-Hodgkin's lymphomas.";
 RL Biochem. Biophys. Res. Commun. 238:273-276(1997).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Neville M.J., Milner C.M., Campbell R.D.,
 "A new member of the immunoglobulin superfamily and a V-ATPase G
 subunit are amongst the predicted products of novel genes close to the
 TNF locus in the human MHC.";
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Rowen L., Madan A., Qin S., Shaffer T., James R., Ratcliffe A.,
 Abbasi N., Dickhoff R., Loretz C., Madan A., Dors M., Young J.,
 Laakey S., Hood L.,
 "Sequence of the human major histocompatibility complex class III
 region.";
 RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RA Shina S., Tamiya G., Oka A., Inoko H.,
 "Homo sapiens 2,229,817bp genomic DNA of 6p21.3 HLA class I region.";
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2), AND VARIANTS GLU-70 AND
 PRO-111.
 RA Rieder M.J., Arnel T.Z., Carrington D.P., Chung M.-W., Lee K.L.,
 Poel C.L., Toth E.J., Yi O., Nickerson D.A.,
 Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Cytokine that binds to LTB/TNFSF3. May play a specific
 role in immune response regulation. Provides the membrane anchor
 for the attachment of the heterotrimeric complex to the cell
 surface. Isoform 2 is probably non-functional.
 CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or
 CC (less prevalent) one LTB and two LTA subunits.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).

CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
 CC produced by alternative splicing.
 CC -1- TISSUE SPECIFICITY: SPLEEN AND THYMUS.
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>
 CC or send an email to license@isb-sdb.ch).
 CC -----

DR EMBL; L11016; AAA9988.1; -;
 DR EMBL; U89922; AAC51769.1; -;
 DR EMBL; U79029; AAB37342.1; -;
 DR EMBL; L11015; AAA36191.1; -;
 DR EMBL; Y14768; CAA75069.1; -;
 DR EMBL; AF129756; AAD18089.1; -;
 DR EMBL; AP000505; BAB63395.1; -;
 DR EMBL; AY070219; AAL49954.1; -;
 DR EMBL; AY070219; AAL49955.1; -;
 DR PIR; A46066; A46066.
 DR HSSP; P01374; 1TNR.
 DR Gene; HGNC:6711; LTB.
 DR MIM; 600978; -;
 DR Interpro: IPR003636; TNF-abc.
 DR Interpro: IPR00478; TNF-family.
 DR Pfam; PF00229; TNF; 1.
 DR PRINTS; PR01234; TNECROSISFCT.
 DR ProDom; PD002012; TNF-abc; 1.
 DR SMART; SM00207; TNF; 1.
 DR PROSITE; PS00251; TNF_1; 1.
 DR PROSITE; PS0049; TNF_2; 1.
 KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor;
 KW Alternative splicing; Polymorphism.
 FT DOMAIN 1 18
 FT TRANSMEM 19 48
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
 FT (POTENTIAL).
 FT DOMAIN 222 222
 FT CAROHD 222 222
 FT VARSPLIC 53 77
 FT VARSPLIC 78 244
 FT VARIANT 70 70
 FT VARIANT 111 111
 FT CONFLICT 60 69
 FT SEQUENCE 244 AA; 25390 MW; F41569459830ED4C CRC64;
 SQ

Query Match 2.5%; Score 7; DB 1; Length 244;
 Best Local Similarity 100.0%; Pred. No. 49;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 194 GLYLYC 200
 DB 131 GLYLYC 137

RESULT 37
 MOEB_ECOLI STANDARD; PRT: 249 AA.

AC P12882;
 DT 01-OCT-1989 (Rel. 12, Created)
 DT 01-OCT-1989 (Rel. 12, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Molycoprotein biosynthesis protein moeb.
 GN MOEB OR CHLN OR B0826.
 OS Escherichia coli.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 CC Escherichia coli.

OX NCBI_TaxId=562;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88314906; PubMed=3045084;
RA Nohno T., Kasai Y., Saito T.;
RT Cloning and sequencing of the *Escherichia coli* chlN operon involved
in molybdopterin biosynthesis.";
RL J. Bacteriol. 170:4097-4102(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE=97426617; PubMed=9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,
RA Mau B., Shao Y.;
RT The complete genome sequence of *Escherichia coli* K-12.";
RL Science 277:1453-1474(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=K12;
RX MEDLINE=97061202; PubMed=8905232;
RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,
RA Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,
RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,
RA Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N.,
RA Sempel G., Seki Y., Tagami H., Takemoto K., Mada C., Yamamoto Y.,
RA Yano M., Horiuchi T.;
RT A 7.18-Kb DNA sequence of the *Escherichia coli* K-12 genome
corresponding to the 12.7-28.0 min region on the linkage map.";
RL DNA Res. 3:137-155(1996).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
(MOLYBDOPTERIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE
ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTING
FACTOR (MOAD).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M21151; AAA23580.1; -;
DR EMBL: AE000185; AAC73913.1; -;
DR EMBL: D90720; BAA35514.1; -;
DR EMBL: D90721; BAA35521.1; -;
DR PIR: B32352; B32352;
DR EcoGene: EGI0154; moeb.
DR InterPro: IPR000205; NAD-binding.
DR InterPro: IPR000594; Thif_domain.
DR Pfam: PF00899; Thif. 1.
KW Molybdenum cofactor biosynthesis; Complete proteome.
SQ SEQUENCE 249 AA; 26719 MW; 12C77082B3F39D7D CRC64;

Query Match 2.5%: Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNRQI 185
Db 10 LRYNRQI 16
|||||

RESULT 38
MOEB_SALTY STANDARD; PRT; 249 AA.
AC 056067;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Molybdopterin biosynthesis protein moeb.
GN MOEB OR STM0845.
OS *Salmonella typhimurium*.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC *Salmonella*.
OX NCBI_TaxId=602;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2;
RA Wong K.K., Kwan H.S.;
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=LT2 / SGSC1412 / ATCC 700720;
RX MEDLINE=21534948; PubMed=11677609;
RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,
RA Courtney L., Porwollik S., Ali J., Dante M., Du F., Hou S., Layman D.,
RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,
RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,
RA Waterston R., Wilson R.K.;
RT "Complete genome sequence of *Salmonella enterica* serovar Typhimurium
LT2.";
RL Nature 413:852-856(2001).
CC -1- FUNCTION: INVOLVED IN BIOSYNTHESIS OF A DEMOLYBDO COFACTOR
(MOLYBDOPTERIN), NECESSARY FOR MOLYBDENZYMES. PLAYS A ROLE IN THE
ACTIVATION OF THE SMALL SUBUNIT OF THE MOLYBDOPTERIN CONVERTING
FACTOR (MOAD).
CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE HESA/MOEB/THIF FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U53178; AAA96530.1; -;
DR EMBL: AE008735; AAL19781.1; -;
DR StGene: SG1065; moeb.
DR InterPro: IPR000205; NAD-binding.
DR InterPro: IPR000594; Thif_domain.
DR Pfam: PF00899; Thif. 1.
KW Molybdenum cofactor biosynthesis; Complete proteome.
FT CONFLICT 9 9 M->I (IN REF. 1).
FT CONFLICT 38 38 G->R (IN REF. 1).
FT CONFLICT 117 117 S->A (IN REF. 1).
FT CONFLICT 169 169 N->T (IN REF. 1).
FT CONFLICT 221 221 G->E (IN REF. 1).
SQ SEQUENCE 249 AA; 26903 MW; 0F0050831D557AD2 CRC64;

Query Match 2.5%: Score 7; DB 1; Length 249;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 179 LRYNRQI 185
Db 10 LRYNRQI 16
|||||

RESULT 39
KDTX_SERMA STANDARD; PRT; 257 AA.
AC 054435;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE Lipopolysaccharide core biosynthesis glycosyl transferase kdtX
(EC 2.-.-.-).
GN KDTX.
OS *Serratia marcescens*.

```

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Serratia.
OX NCBI_TaxID=615;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=N28B;
RX MEDLINE=96422003; PubMed=8824620;
RA Guasch J.F., Pique N., Climent N., Ferrer S., Merlino S., Rubires X.,
RA Tomas J.M., Regue M.;
RT "Cloning and characterization of two Serratia marcescens genes
RT involved in core lipopolysaccharide biosynthesis.";
RL J. Bacteriol. 178:5741-5747(1996)
CC -1- PATHWAY: Lipopolysaccharide core biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE GLYCOSTYLTRANSFERASE FAMILY 2. WAE/KDTX
CC SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U52844; AAC4443.1; -
DR InterPro: IPR001173; Glycosyltransf_2.
DR Pfam: PF00535; Glycosyltransf_2; 1.
DR Lipopolysaccharide biosynthesis; Transferase; Glycosyltransferase.
SQ SEQUENCE 257 AA; 29233 MW; D40D7B57E002F990 CRC64;

Query Match 2.5%; Score 7; DB 1; Length 257;
Best Local Similarity 100.0%; Pred. No. 51;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 71 GLLAVV 77
|111111|
DB 229 GLLAVV 235

RESULT 40
TNFC_MARMO
ID TNFC_MARMO STANDARD; PRT; 310 AA.
AC 09JUN10: 09JUN11;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lymphotoxin-beta (LT-beta) (Tumor necrosis factor C) (TNF-C) (Tumor
DE necrosis factor ligand superfamily member 3).
GN LTB OR TNFSF3 OR TNFC.
OS Marmota monax (Woodchuck).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Scuridae; Sciurinae;
OC Marmota.
OX NCBI_TaxID=9995;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20184748; PubMed=10721723;
RA Li D.H., Haveli E.A., Brown C.L., Cullen J.M.;
RT "Woodchuck lymphotoxin-alpha, -beta and tumor necrosis factor genes:
RT structure, characterization and biological activity.";
RL Gene 242:295-305(2000).
CC -1- FUNCTION: Cytokine that binds to LTB/TNFSF3. May play a specific
CC role in immune response regulation. Provides the membrane anchor
CC for the attachment of the heterotrimeric complex to the cell
CC surface.
CC -1- SUBUNIT: Heterotrimer of either two LTB and one LTA subunits or
CC (less prevalent) two LTA and one LTB subunits (By similarity).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein (Potential).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration -
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF096268; AAF3486.1; -
DR EMBL: AF095587; AAF3486.1; -
DR HSSP: P01374; 1TNR.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF; 1.
DR PRINTS: PR01234; TNFCROSISFCT.
DR PRODOM: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS50049; TNF_2; 1.
DR CYTOKINE: Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor.
FT DOMAIN 1 27 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 28 48 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT (POTENTIAL).
FT DOMAIN 49 310 EXTRACELLULAR (POTENTIAL).
FT CARBOHYD 272 272 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 280 280 D -> H (IN REF. 1; AAF34865).
SQ SEQUENCE 310 AA; 32644 MW; 73B354EFC8B3B3BE CRC64;

Query Match 2.5%; Score 7; DB 1; Length 310;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 194 GLYLYYC 200
|111111|
DB 181 GLYLYYC 187

```

Search completed: March 31, 2003, 14:14:17
Job time : 30 secs

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 13:51:31 ; Search time 76 Seconds
(without alignments)
769,966 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 284
Sequence: 1 MSULDFEISARLRPLRSLG.....PNAHLKAPFLTYFGLFQVH 284

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 65 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	10	3.5	111	16	08X4J8
2	9	3.2	142	16	0981J0
3	8	2.8	51	4	09BX52
4	8	2.8	143	17	09HST7
5	8	2.8	151	10	09SD11
6	8	2.8	184	5	09VV70
7	8	2.8	197	16	09WYU0
8	8	2.8	199	4	09BU11
9	8	2.8	211	5	08SXH4
10	8	2.8	278	16	09ZG99
11	8	2.8	279	11	09D378
12	8	2.8	279	11	09CPR8
13	8	2.8	306	5	08SZB8
14	8	2.8	339	17	058554
15	8	2.8	342	4	099908
16	8	2.8	343	4	09BPV2

ALIGNMENTS

RESULT 1	ID	08X4J8	PRELIMINARY:	PRT:	111 AA.
AC	08X4J8	01-MAR-2002 (TREMBlrel. 20, Created)			
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)				
DT	01-JUN-2002 (TREMBlrel. 21, Last annotation update)				
DE	Hypothetical protein 23516.				
GN	23516.				
OS	Escherichia coli O157:H7.				
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
OC	Escherichia.				
OX	NCBI_TaxID=83334;				
RN	11				
RP	SEQUENCE FROM N.A.				
RC	SPTRAIN-O157:H7 / EDL933 / ATCC 700927;				
RX	MEDLINE-21074935; PubMed-11206551;				
RA	Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,				

17	8	2.8	370	17	08TU08
18	8	2.8	372	16	053860
19	8	2.8	377	16	09KRT3
20	8	2.8	387	10	09SHD8
21	8	2.8	397	16	080820
22	8	2.8	435	16	08XS70
23	8	2.8	443	4	09UFM5
24	8	2.8	444	5	09VNP0
25	8	2.8	455	16	08ZCV8
26	8	2.8	471	16	09RR18
27	8	2.8	472	4	096N66
28	8	2.8	473	11	09CV76
29	8	2.8	473	11	08R1P9
30	8	2.8	522	10	09L1W0
31	8	2.8	564	12	09YT09
32	8	2.8	568	12	08QPL0
33	8	2.8	656	5	09N8H2
34	8	2.8	1523	8	09N808
35	7	2.5	24	5	P83215
36	7	2.5	30	5	P83217
37	7	2.5	32	6	09BER6
38	7	2.5	35	2	09ZG35
39	7	2.5	54	4	016193
40	7	2.5	67	2	08ROS4
41	7	2.5	69	5	09VMM9
42	7	2.5	86	16	09AAG4
43	7	2.5	94	4	09P074
44	7	2.5	99	2	P97199
45	7	2.5	100	5	08WQR8
46	7	2.5	100	12	08QRM6
47	7	2.5	101	15	08UTL1
48	7	2.5	108	9	038214
49	7	2.5	115	16	09HY60
50	7	2.5	118	2	0939P7
51	7	2.5	118	16	0928G5
52	7	2.5	119	10	094LX6
53	7	2.5	121	10	08WOM1
54	7	2.5	123	16	08UCH2
55	7	2.5	125	5	09UJP6
56	7	2.5	126	1	0977X7
57	7	2.5	133	15	091121
58	7	2.5	135	4	09BTR9
59	7	2.5	145	4	09UMX4
60	7	2.5	150	4	09Y413
61	7	2.5	156	10	09ACR8
62	7	2.5	157	10	049178
63	7	2.5	157	16	09RYB7
64	7	2.5	157	16	09RXG1
65	7	2.5	159	10	08WOT0

08TU08	methanopyru
053860	mycobacteri
09KRT3	streptomyc
09SHD8	arabidopsis
080820	agrobacteri
08XS70	ralstonia s
09UFM5	homo sapien
09VNP0	drosophila
08ZCV8	yeast
09RR18	delnoco
096N66	homo sapien
09CV76	mus musculu
08R1P9	mus musculu
09L1W0	oryza sativ
09YT09	influenzavi
08QPL0	influenza a
09N8H2	trypanosoma
09N808	trypanosoma
P83215	octopus vul
P83217	octopus vul
09BER6	capra hircu
09ZG35	chlamydia t
016193	homo sapien
08ROS4	seriatia ma
09VMM9	drosophila
09AAG4	caulobacter
09P074	homo sapien
P97199	escheric
08WQR8	eupryma sc
08QRM6	hepatitis c
08UTL1	human immun
038214	bacterioph
09HY60	pseudomon
0939P7	pseudomon
0928G5	listeria in
094LX6	euphorbia t
08WOM1	oryza sativ
08UCH2	agrobacteri
09UJP6	caenorhabd
0977X7	methanococ
091121	human immun
09BTR9	homo sapien
09UMX4	homo sapien
09Y413	homo sapien
09ACR8	oryza sativ
049178	oryza sativ
09RYB7	delnoco
09RXG1	delnoco
08WOT0	sorghum btl

RA Posfai G., Hackett J., Klink S., Boufin A., Shao Y., Miller L.,
 RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamouzis K.,
 RA Apodaca J., Anantharaman T.S., Lin J., Ten G., Schwartz D.C.,
 RA Welch R.A., Blatner F.R.,
 RT "Genome sequence of enterohaemorrhagic *Escherichia coli* O157:H7.";
 RL Nature 409:529-533(2001).
 DR EMBL: AE005458; AAG57389.1; -.
 DR InterPro: IPR00620; DUF6.
 DR Pfam: PF00892; DUF6; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 111 AA; 12165 MW; 7CEFC93D786CD759 CRC64;

Query Match 3.5%; Score 10; DB 16; Length 111;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGIALACGL 72
 DB 40 LGIALACGL 49

RESULT 2

O981J0 PRELIMINARY; PRT; 142 AA.
 ID O981J0:
 AC O981J0: 01-OCT-2001 (TREMBlrel. 18, Created)
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
 DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
 DE Probable transport protein.
 GN MLR2380.
 OS Rhizobium loti (Mesorhizobium loti).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Phyllobacteriaceae; Mesorhizobium.
 OX NCBI_TaxID=381;

RA [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MAEF303099;
 RX MEDLINE=21082930; PubMed=11214968;
 RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
 RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,
 RA Wochizuki Y., Nakayama S., Nakazaki N., Shimo S., Sugimoto M.,
 RA Takeuchi C., Yamada M., Tabata S.;
 RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
 RT Mesorhizobium loti.";
 RL DNA Res. 7:331-338(2000).
 DR EMBL: AP002999; BAB49526.1; -.
 KW Complete proteome.
 SQ SEQUENCE 142 AA; 14884 MW; 0DCA7842CB5A5E6F CRC64;

Query Match 3.2%; Score 9; DB 16; Length 142;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 53 TALVPLAL 61
 DB 122 TALVPLAL 130

RESULT 3

O9BX52 PRELIMINARY; PRT; 51 AA.
 ID O9BX52:
 AC O9BX52: 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE DJ81IC21.1 (Novel protein similar to gamma-glutamyl
 DE transpeptidase-related protein (GGT-rel)) (Fragment).
 GN DJ81IC21.1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;

RN [1]
 RP SEQUENCE FROM N.A.
 RA Corby N.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AL133466; CAC34607.1; -.
 FT NON_TER 51 51
 SQ SEQUENCE 51 AA; 5124 MW; 31980CCA0451E8F CRC64;

Query Match 2.8%; Score 8; DB 4; Length 51;
 Best Local Similarity 100.0%; Pred. No. 5.8;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68
 DB 7 LGIGLALA 14

RESULT 4

O9HST7 PRELIMINARY; PRT; 143 AA.
 ID O9HST7:
 AC O9HST7: 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)
 DE Vn90080H.
 GN Vn90080H.
 OS Halobacterium sp. (Strain NRC-1).
 OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
 OC Halobacteriaceae; Halobacterium.
 OX NCBI_TaxID=64091;
 RN [1]

RP SEQUENCE FROM N.A.
 RX MEDLINE=20504483; PubMed=11016950.
 RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,
 RA Shukla H.D., Lasky S.R., Balliga N.S., Thorsson V., Shrogha J.,
 RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
 RA Lelhauser B., Keller K., Cruz R., Danon M.J., Hough D.W.,
 RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angelino C.M., Dale H.,
 RA Isehaber T.A., Beck R.F., Pohlshroder M., Spidlich J.L., Jung K.-H.,
 RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
 RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
 RT "Genome sequence of Halobacterium species NRC-1.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
 DR EMBL: AE004976; AAG18715.1; -.
 KW Complete proteome.
 SQ SEQUENCE 143 AA; 15648 MW; 45466E328EF3468 CRC64;

Query Match 2.8%; Score 8; DB 17; Length 143;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGLALA 68
 DB 55 LGIGLALA 62

RESULT 5

O9SD11 PRELIMINARY; PRT; 151 AA.
 ID O9SD11:
 AC O9SD11: 01-MAY-2000 (TREMBlrel. 13, Created)
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Hypothetical protein (OSJNB40036E02.6 protein) (B1085F09.2
 DE protein).
 GN B1085F09.2.
 OS Oryza sativa (Rice).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Eurnatoidae; Oryzaceae; Oryza.
 OX NCBI_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN-CV, NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, PAC
 clone: P000310.";
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, BAC
 clone: OSJNB80036E02.";
 RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-CV, NIPPONBARE;
 RA Sasaki T., Matsumoto T., Yamamoto K.;
 RT "Oryza sativa nippobare(GA3) genomic DNA, chromosome 1, BAC
 clone: B1085F09.";
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF000815; BAA87834.1; -;
 DR EMBL: AF002862; BAB1732.1; -;
 DR EMBL: AF003103; BAB44106.1; -;
 SQ SEQUENCE 151 AA; 16632 MW; EC68451ECA2BD71D CRC64;

Query Match 2.8%; Score 8; DB 10; Length 151;
 Best Local Similarity 100.0%; Pred. No. 15;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 43 RRRGRGGE 50
 Db 131 RRRGRGGE 138

RESULT 6

ID Q9VV70 PRELIMINARY; PRT; 184 AA.
 AC Q9VV70;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE CG13033 protein.
 GN CG13033.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Braachyera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazel R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cwley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland J.J., Wei M.-H., Idegawa C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lai Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Sytkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003527; AAF49452.1; -;
 DR FLYBase: FBgn0036638; CG13033.
 DR InterPro: IPR004019; YLP_motif.
 DR Pfam: PF02757; YLP_5.
 SQ SEQUENCE 184 AA; 21022 MW; 31976AE350DDC447 CRC64;

Query Match 2.8%; Score 8; DB 5; Length 184;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 72 LLAAVSL 79
 Db 16 LLAAVSL 23

RESULT 7

ID Q9WYU0 PRELIMINARY; PRT; 197 AA.
 AC Q9WYU0;
 DT 01-NOV-1999 (TREMBLrel. 12, Created)
 DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical protein TM0469.
 GN TM0469.
 OS Thermotoga maritima.
 OC Bacteria; Thermotogae; Thermotogales; Thermotogaceae; Thermotoga.
 OX NCBI_TaxID=2336;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-MSB / DSM 3109;
 RX MEDLINE=99287316; PubMed=10360571;
 RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,
 RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
 RA McDonald L., Uitterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
 RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,
 RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
 RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
 RT "Evidence for lateral gene transfer between Archaea and Bacteria from
 genome sequence of Thermotoga maritima.";
 RL Nature 399:323-329(1999).
 DR EMBL: AE001724; AAD35553.1; -;
 DR TIGR: TM0469;
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 197 AA; 22919 MW; 41E2C8E3C09180BC CRC64;

Query Match 2.8%; Score 8; DB 16; Length 197;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 2 SLDPFEIS 9
 Db 135 SLDPFEIS 142

RESULT 8
 Q9BU11 PRELIMINARY; PRT; 199 AA.
 ID Q9BU11

```

AC 09BUI1;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 22.8 kda protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
OX NCBI_TaxID=9606;
RN 11;
RP SEQUENCE FROM N.A.
RC TISSUE-BRAIN;
RA Strausberg R.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC002598; AAH02598.1;
DR InterPro; IPR004299; MBOAT_fam.
DR Pfam; PF03062; MBOAT; 1.
KM Hypothetical protein.
SQ SEQUENCE 199 AA; 22750 MW; E5DACT47E7E8F06B CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 4; Length 199;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALGLAL 67
DB 167 ALGLAL 174
|||||||

RESULT 9
OBSX4 PRELIMINARY; PRT; 211 AA.
ID 08SX4;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE RES0345D.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RA Stapleton M., Brokslein P., Hong L., Agbayani A., Carlson J.,
RA Champagne M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rudin G.M.,
RA Ceiniker S.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY089634; AAL90372.1;
SQ SEQUENCE 211 AA; 23780 MW; 82FF4983E91F510A CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 5; Length 211;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVSL 79
DB 9 LLLAVSL 16
|||||||

RESULT 10
O9ZG99 PRELIMINARY; PRT; 278 AA.
ID 09ZG99;
AC 09ZG99;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Transmembrane protein AMPE.
RA AMPE OR PA4521.

```

```

OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC Pseudomonas.
OX NCBI_TaxID=287;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RA Langae Y.T., Dargis M., Huletsky A.;
RT "An ampD gene in Pseudomonas aeruginosa encodes a negative regulator
of ampC beta-lactamase expression.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN 12;
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043;
RA Stoyer C.K., Pham X.-O.T., Ervin A.L., Mizoguchi S.D., Warren P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Coltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
opportunistic pathogen.";
RL Nature 406:959-964(2000).
DR EMBL; AF082575; AAC96784.1;
DR EMBL; AE004866; AAG07909.1;
KW Transmembrane; Complete proteome.
SQ SEQUENCE 278 AA; 30793 MW; C623F1AB0691CFEE CRC64;

Query Match
Best Local Similarity 100.0%; Score 8; DB 16; Length 278;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 70 LGLLAV 77
DB 47 LGLLAV 54
|||||||

RESULT 11
O9D378 PRELIMINARY; PRT; 279 AA.
ID 09D378;
AC 09D378;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 5730494616Rik protein.
GN 5730494616Rik.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN 11;
RP SEQUENCE FROM N.A.
RC STRAIN-C57BL/6J; TISSUE-MEDULLA OBLONGATA;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Aizawa T., Hara A., Fukunishi Y., Kono H., Aachii J., Fukuda S.,
RA Arakawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochia H.,
RA Kuwaji P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,
RA Schriml L.M., Stabili F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyokawa K., Wang K.H., Weitz C., Wittaker C., Wilming L.,
RA Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;

```

RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK018250; BAB31133.1; -
 DR MGD: MGI:1913897; 5730494G16RIK.
 DR InterPro: IPR002190; MAGE.
 DR Pfam: PF01454; MAGE.1.
 SQ SEQUENCE 279 AA; 31474 MW; 5E243590A99F15F0 CRC64;

Query Match 2.8%; Score 8; DB 11; Length 279;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 GSRASLSA 87
 Db 41 GSRASLSA 48

RESULT 12
 Q9CPR8 PRELIMINARY; PRT; 279 AA.

AC Q9CPR8;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
 DE 5730494G16RIK protein (MAGE-91).
 GN 5730494G16RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]

RP SEQUENCE FROM N.A.
 RC STRAIN=CS7BL/6J; TISSUE=EMBRYO, AND EMBRYONIC STEM CELLS;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Atakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Pleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Ginstacinc S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
 RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Augier P.H., Chomez P.M., De Backer O.R., Bertrand M.J.M.;
 RT "Ten new murine members of the MAGE gene family."
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK017727; BAB30899.1; -
 DR EMBL: AK010294; BAB26830.1; -
 DR EMBL: AF319979; AAK01207.1; -
 DR MGD: MGI:1913897; 5730494G16RIK.
 DR InterPro: IPR002190; MAGE.
 DR Pfam: PF01454; MAGE.1.
 SQ SEQUENCE 279 AA; 31460 MW; FE2435919B63160 CRC64;

Query Match 2.8%; Score 8; DB 11; Length 279;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 80 GSRASLSA 87
 Db 41 GSRASLSA 48

Db 41 GSRASLSA 48

RESULT 13

AC Q8S2B8 PRELIMINARY; PRT; 306 AA.
 DT 01-JUN-2002 (TReMBLrel. 21, Created)
 DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
 DE 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE R07882P.
 GN CG1169.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Abhayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
 RA George R., Gonzalez M., Guarin H., Krommiller B., Li P., Liao G.,
 RA Miranda A., Mungall C.J., Nunoo J., Paclebo J., Paragas V., Park S.,
 RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
 RA Ceiniker S.;

RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY070982; AAL48604.1; -
 SQ SEQUENCE 306 AA; 34083 MW; 32B69371475448F9 CRC64;

Query Match 2.8%; Score 8; DB 17; Length 306;
 Best Local Similarity 100.0%; Pred. No. 29;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 56 LVPALGL 63
 Db 170 LVPALGL 177

RESULT 14
 ID 058554 PRELIMINARY; PRT; 339 AA.
 AC 058554;
 DT 01-AUG-1998 (TReMBLrel. 07, Created)
 DT 01-JAN-1999 (TReMBLrel. 09, Last sequence update)
 DT 01-MAR-2002 (TReMBLrel. 20, Last annotation update)
 DE Hypothetical protein PH0824.
 GN PH0824.
 OS Pyrococcus horikoshii.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=53953;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OT3;
 RX MEDLINE=98344137; PubMed=9679194;
 RA Kawarabayashi Y., Sawada M., Horikawa H., Halkawa Y., Hino Y.,
 RA Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,
 RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohkura Y.,
 RA Funahashi T., Tanaka T., Kudo H., Yamazaki J., Kishida N., Oguchi A.,
 RA Aoki K.-I., Yoshizawa T., Nakamura Y., Robb F.T., Horikoshi K.,
 RA Masuchi Y., Shizuya H., Kikuchi H.;

RT "Complete sequence and gene organization of the genome of a hyper-thermophilic archaeobacterium, Pyrococcus horikoshii OT3."
 RL DNA Res. 5:55-76(1998).
 DR EMBL: AP000003; BAA29917.1; -
 DR InterPro: IPR002657; Billeac/Na_smprtr.
 DR Pfam: PF01758; SBF.1.
 KW Hypothetical protein, Complete proteome

Query Match 2.8%; Score 8; DB 17; Length 339;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 70 LGLLAVV 77
 |||||
 DB 111 LGLLAVV 118

RESULT 15
 ID 099908 PRELIMINARY; PRT: 342 AA.
 AC 099908:
 DT 01-MAY-1997 (TREMBLrel. 03, Created)
 DT 01-MAY-1997 (TREMBLrel. 03, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE BBI protein.
 GN BBI.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE-96273128; PubMed-8702217;
 RX Fukushima-Johnson N., Lee S.W., Liebert M., Grossman H.B.;
 RT "Molecular analysis of a gene, BBI, overexpressed in bladder and
 RT breast carcinoma."
 RL Anticancer Res. 16:1085-1090(1996).
 DR EMBL: S82470; AAB37433.1; -;
 DR InterPro: IPR004299; MBOAT_fam.
 DR Pfam: PF03062; MBOAT_1.
 SQ SEQUENCE 342 AA; 38163 MW; 2B479E8CF1B91C CRC64;

Query Match 2.8%; Score 8; DB 4; Length 342;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 ALGGLAL 67
 |||||
 DB 310 ALGGLAL 317

RESULT 16
 ID 09BPV2 PRELIMINARY; PRT: 343 AA.
 AC 09BPV2:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Unknown (protein for MGC:4221) (protein for MGC:2029).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY, AND COLON;
 RA Strusberg R.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC003164; AAB03164.1; -;
 DR EMBL: BC002513; AAB02512.1; -;
 DR InterPro: IPR004299; MBOAT_fam.
 DR Pfam: PF03062; MBOAT_1.
 SQ SEQUENCE 343 AA; 38727 MW; F71E7DBF7ABD9BB7 CRC64;

Query Match 2.8%; Score 8; DB 4; Length 343;
 Best Local Similarity 100.0%; Pred. No. 32;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 ALGGLAL 67
 |||||
 DB 311 ALGGLAL 318

RESULT 17
 ID 08TU8 PRELIMINARY; PRT: 370 AA.
 AC 08TU8:
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Permease subunit of a ABC-type transport system involved in
 DE lipoprotein release.
 GN MK1655.
 OS Methanopyrus kandleri.
 OC Archaea; Euryarchaeota; Methanopyri; Methanopyrales; Methanopyraceae;
 OC Methanopyrus.
 OX NCBI_TaxID=2320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AV19 / DSM 6324 / JCM 9639;
 RX MEDLINE-21927647; PubMed-11930014;
 RA Slesarev A.I., Mezhevaya K.V., Makarova K.S., Polushin N.N.,
 RA Shcherbina O.V., Shakhova V.V., Belova G.I., Aravind L.,
 RA Natale D.A., Rogozin I.B., Talusov R.L., Wolf Y.I., Stetter K.O.,
 RA Mal'kh A.G., Koonin E.V., Kozlovskiy S.A.;
 RT "The complete genome of hyperthermophile Methanopyrus kandleri AV19
 RT and monophyly of archaeal methanogens."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:4644-4649(2002).
 DR EMBL: AE010455; AAM02868.1; -;
 KW Complete proteome.
 SQ SEQUENCE 370 AA; 39411 MW; B07662EAE1E5A644E CRC64;

Query Match 2.8%; Score 8; DB 17; Length 370;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 61 LGGLALA 68
 |||||
 DB 336 LGGLALA 343

RESULT 18
 ID 053860 PRELIMINARY; PRT: 372 AA.
 AC 053860:
 DT 01-JUN-1998 (TREMBLrel. 06, Created)
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein cyam3.
 GN CYM3 OR RV0848 OR MTV043.41.
 OS Mycobacterium tuberculosis.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE-98295987; PubMed-9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham D., Gentles S., Hamlin N., Holroyd S.,
 RA Horsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence."
 RL Nature 393:537-544(1998).
 DR EMBL: AL022004; CAAT1654.1; -;
 DR HSP: P35520; LJBQ.
 DR Tuberculin: RV0848; -;
 DR InterPro: IPR001926; B6_enzyme_beta.
 DR Pfam: PF00291; PALP; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 372 AA; 40118 MW; 927386BE1DF5FB6C CRC64;

Query Match 2.8%; Score 8; DB 16; Length 372;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LAGLALA 68
 |||||
 Db 96 LAGLALA 103

RESULT 19
 O9KTY3

PRELIMINARY; PRT; 377 AA.

AC O9KTY3; PRT; 377 AA.
 DT 01-OCT-2000 (TREMUREL. 15, Created)
 DT 01-OCT-2000 (TREMUREL. 15, Last sequence update)
 DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
 DE Putative integral membrane protein.
 GN SC05682 OR SC5H4.06C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;

RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdano-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrall B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RL Nature 417:141-147(2002).
 DR EMBL; AL355913; CAB91118.1; -
 SO SEQUENCE 377 AA; 37614 MW; A35DA0437F044C6 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 377;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 59 LAGLALA 66
 |||||
 Db 327 LAGLALA 334

RESULT 20
 O9SHD8

PRELIMINARY; PRT; 387 AA.

AC O9SHD8; PRT; 387 AA.
 DT 01-MAY-2000 (TREMUREL. 13, Created)
 DT 01-MAY-2000 (TREMUREL. 13, Last sequence update)
 DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
 DE At2g45000 protein (At2g45000/T14P1.20) (Hypothetical 40.6 kDa
 protein).
 GN At2g45000 OR At2g45005.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;

RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA MEDLINE=20083487; Pubmed=10617197;
 RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
 RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblum T.V.,
 RA Beil C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
 RA Cronin L.A., Shen M., Vanaken S.E., Umayam L., Tallon L.J., Gail J.E.,

RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
 RA Copenhagen G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
 RA Salzberg S.L., Fraser C.M., Venter J.C.;
 RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
 thaliana."
 RL Nature 402:761-768(1999).

RP SEQUENCE FROM N.A.
 RC STRAIN-CV. COLUMBIA;
 RA Lin X.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]

RP SEQUENCE FROM N.A.
 RA Shinn P., Chen H., Cheuk R., Kim C.J., Lim J., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RT "Arabidopsis cDNA clones."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]

RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Arabidopsis full length cDNA clones."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AC007659; RAD32835.1; -
 DR EMBL; AY074646; AAL69462.1; -
 DR EMBL; AY080620; AAL69303.1; -
 KW Hypothetical protein.
 SO SEQUENCE 387 AA; 40584 MW; AF6C6B3BAC9BF69A CRC64;

Query Match 2.8%; Score 8; DB 10; Length 387;
 Best Local Similarity 100.0%; Pred. No. 35;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 EELVAED 100
 |||||
 Db 370 EELVAED 377

RESULT 21
 O8U820

PRELIMINARY; PRT; 397 AA.

AC O8U820; PRT; 397 AA.
 DT 01-JUN-2002 (TREMUREL. 21, Created)
 DT 01-JUN-2002 (TREMUREL. 21, Last sequence update)
 DT 01-JUN-2002 (TREMUREL. 21, Last annotation update)
 DE Hypothetical protein AtU3948.

GN AtU3948 OR AGR.L.1808.
 OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
 OC Rhizobiaceae; Rhizobium.
 OX NCBI_TaxID=176299;

RP SEQUENCE FROM N.A.
 RA MEDLINE=21608550; Pubmed=11743193;
 RA Wood D.W., Setubal J.C., Kaul R., Monks D.E., Kitajima J.P.,
 RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
 RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D., Sr.,
 RA Chapman P., Clendenning J., Deatherage G., Gillet W., Grant C.,
 RA Kutayavon T., Levy R., Li M.-J., McClelland E., Palmeri A.,
 RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
 RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
 RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,

RA Chunmley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
 RA Neater E.W.;
 RT "The genome of the natural genetic engineer *Agrobacterium tumefaciens*
 RT C58";
 RL Science 294:2317-2323(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21608551; PubMed=11743194;
 RA Goodner B., Hinkle G., Gatlung S., Miller N., Blanchard M.,
 RA Guoillo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
 RA Houmlet K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
 RA Flanagan C., Allinger M., Doughy D., Scott C., Iappas C., Markelz B.,
 RA Flanagan C., Crowell C., Gursion J., Lomo C., Sear C., Strub G.,
 RA Cielo C., Slater S.;
 RT "Genome sequence of the plant pathogen and biotechnology agent
 RT *Agrobacterium tumefaciens* C58.";
 RL Science 294:2323-2328(2001).
 DR EMBL: AE009325; AAL44750.1; ALT_INIT.
 DR EMBL: AE008289; AAK89478.1; -.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 397 AA; 41708 MW; 700748E32A46AE86 CRC64;

Query Match 2.8%; Score 8; DB 16; Length 397;
 Best Local Similarity 100.0%; Pred. No. 36;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 251 LALRPGSS 258
 DB 341 LALRPGSS 348

RESULT 22

ID Q8XS70 PRELIMINARY; PRT; 435 AA.
 AC Q8XS70;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
 DE Probable transmembrane protein.
 GN RSP0611 OR RS03756.
 OS *Ralstonia solanacearum* (Pseudomonas solanacearum).
 OC Plasmid megaplasmid.
 OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
 OC Ralstonia.
 OX NCBI_TaxID=305;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=GM11000;
 RX MEDLINE=21681879; PubMed=11823852;
 RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
 RA Arlat M., Billault A., Brottier P., Camus J.C., Cattolico L.,
 RA Chaudler M., Choisme N., Claudel-Renard C., Cunnac S., Demange N.,
 RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schlex T.,
 RA Sigler P., Trebault P., Whalen M., Wincker P., Levy M.,
 RA Weisenbach J., Boucher C.A.;
 RT "Genome sequence of the plant pathogen *Ralstonia solanacearum*";
 RL Nature 415:497-502(2002).
 DR EMBL: AL646079; CAD17762.1; -.
 DR InterPro: IPR001564; NDK.
 DR PROSITE: PS00469; NDP_KINASES; UNKNOMN_1.
 KW Plasmid; Complete proteome.
 SQ SEQUENCE 435 AA; 47048 MW; CCD859D9C54DDB5A CRC64;

Query Match 2.8%; Score 8; DB 16; Length 435;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 217 LVDGVIAL 224
 DB 358 LVDGVIAL 365

RESULT 23

O9UFM5
 ID O9UFM5 PRELIMINARY; PRT; 443 AA.
 AC O9UFM5;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE Hypothetical 47.4 kDa protein.
 GN DKFZP5660011.
 OS *Homo sapiens* (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=KIDNEY;
 RA Blum H., Baerends S., Mewes H.W., Gassenhuber J., Wiemann S.;
 RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AL117414; CAB55910.1; -.
 DR MEROPS: T03.002; -.
 DR InterPro: IPR00101; G4lutrasptase.
 DR Pfam: PF01019; G_glu_transptase. 1.
 DR PRINTS: PR01210; GGTANSPTASE.
 KW Hypothetical protein.
 SQ SEQUENCE 443 AA; 47446 MW; 3D75E9DB0B265971 CRC64;

Query Match 2.8%; Score 8; DB 4; Length 443;
 Best Local Similarity 100.0%; Pred. No. 40;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 61 LGIGIALA 68
 DB 14 LGIGIALA 21

RESULT 24

ID O9VNP0 PRELIMINARY; PRT; 454 AA.
 AC O9VNP0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
 DE CG1169 protein.
 GN CG1169.
 OS *Drosophila melanogaster* (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champé M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Brill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferrieres S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,
 RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jaitani M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merklov G., Mlshina N.V., Mobery C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Rembert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Slier E., Spreading A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-P., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Glibe R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003600; AAF51889.1; -;
 DR FlyBase: FBgn0037428; CG1169.
 SQ SEQUENCE 454 AA; 51320 MW; A75AAD97E716573 CRC64;

Query Match
 Best Local Similarity 2.8%; Score 8; DB 5; Length 454;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 56 LVPLATGL 63
 Db 170 LVPLATGL 177

RESULT 25
 082CV8
 ID 082CV8 PRELIMINARY; PRT; 465 AA.
 AC 082CV8;
 DT 01-MAR-2002 (TReMBLrel. 20, Created)
 DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE Putative membrane protein.
 GN YEGB OR YPO2850.
 OS *Yersinia pestis*.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC *Yersinia*.
 OX NCBI_TaxID=632;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CO-92 / BIOVAR ORIENTALIS;
 RX MDLINE=21470413; PubMed=11586360;
 RA Parkhill J., Wren B.W., Thomson N.R., Titball R.W., Holden M.T.G.,
 RA Prentice M.B., Sebahia M., James K.D., Churcher C., Mungall K.L.,
 RA Baker S., Basham D., Bentley S.D., Brooks K., Cerdeno-Tarraga A.M.,
 RA Chillingworth T., Cronin A., Davies R.M., Davis P., Dougan G.,
 RA Feltham T., Hamlin N., Holtroyd S., Jagers K., Karlyshev A.V.,
 RA Leather S., Moule S., Oyston P.C.F., Quail M., Rutherford K.,
 RA Simmonds M., Skellon J., Stevens K., Whitehead S., Barrett B.G.;
 RT "Genome sequence of *Yersinia pestis*, the causative agent of plague.";
 RL Nature 413:523-527(2001).
 DR EMBL: A414154; GAC92102.1; -;
 DR InterPro: IPR003662; sub_transporter.
 DR Pfam: PF00083; sugar_tr; 1.
 KW Hypothetical protein; Complete proteome.
 SQ SEQUENCE 465 AA; 50176 MW; 0CC273F10B3F5BD CRC64;

Query Match
 Best Local Similarity 2.8%; Score 8; DB 16; Length 465;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 72 LLAVALSL 79
 Db 335 LLAVALSL 342

RESULT 26
 09RR18

ID 09RR18 PRELIMINARY; PRT; 471 AA.
 AC 09RR18;
 DT 01-MAY-2000 (TReMBLrel. 13, Created)
 DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)
 DE Transpor protein, putative.
 GN DR2502.
 OS *Deinococcus radiodurans*.
 OC Bacteria; Thermus/Deinococcus group; Deinococci; Deinococcales;
 OC Deinococcaceae; Deinococcus.
 OX NCBI_TaxID=1299;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=RI;
 RX MDLINE=20036896; PubMed=10567266;
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
 RA Dodson R.J., Halt D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
 RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
 RA Vamathevan J.J., Lam P., McDonald L., Uterback T., Zaleski C.,
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
 RA Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,
 RA Fraser C.M.;
 RT "Genome sequence of the radioresistant bacterium *Deinococcus*
 RT *radiodurans* RI.";
 RL Science 286:1571-1577(1999).
 DR EMBL: AE002079; AAF12043.1; -;
 DR TIGR: DR2502; -;
 KW Complete proteome.
 SQ SEQUENCE 471 AA; 47974 MW; 96B2BEBF6E45D27 CRC64;

Query Match
 Best Local Similarity 2.8%; Score 8; DB 16; Length 471;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 60 ALGGLAL 67
 Db 366 ALGGLAL 373

RESULT 27
 096N66
 ID 096N66 PRELIMINARY; PRT; 472 AA.
 AC 096N66;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, Last annotation update)
 DE CNA FLJ31346 fls, clone MESAN1000180, highly similar to BBI-malignant
 DE cell expression-enhanced gene/tumor progression-enhanced gene.
 OS *Homo sapiens* (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tashiro H., Yamazaki M., Watanabe K., Kumagai A., Itakura S.,
 RA Fukuzumi Y., Fujimori Y., Komiyama M., Sugiyama T., Irie R.,
 RA Otsuki T., Sato H., Wakamatsu A., Ishi S., Yamamoto J., Isono Y.,
 RA Kawal-Hio Y., Saito K., Nishikawa T., Kimura K., Yamashita H.,
 RA Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K., Wagatsuma M.,
 RA Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,
 RA Kawakami B., Suzuki Y., Sugano S., Nagahari K., Masuno Y., Nagai K.,
 RA Isegaki T.;
 RT "NEDO human cDNA sequencing project.";
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK055908; BAB71043.1; -;
 DR InterPro: IPR004299; MBOAT_fam.
 DR Pfam: PF03062; MBOAT; 1.
 SQ SEQUENCE 472 AA; 52774 MW; EA721998043F9EED CRC64;

Query Match
 Best Local Similarity 2.8%; Score 8; DB 4; Length 472;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 60 ALACGLAL 67
 |||||||
 DB 440 ALACGLAL 447

RESULT 28

ID O9CY76 PRELIMINARY: PRT: 473 AA.
 AC O9CY76:
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE 5730589L02RIK protein;
 GN 5730589L02RIK.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=EMBRYO;
 RX MEDLINE=21085560; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishi Y.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Alkawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Glasl C., King B., Kochiya H.,
 Kuhl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 Schirml L.M., Staudl F., Suzuki R., Tomita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Balarelli R., Barsh G.,
 Blake J., Boileau D., Bojunga N., Carninci P., de Bonaldo M.F.,
 Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 Lyons P., Marchionni L., Mazzarelli J., Momberts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyo-oka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,
 Wynshaw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohlsuk S.,
 Hayashizaki Y.;
 RA "Functional annotation of a full-length mouse cDNA collection";
 RT Nature 409:685-690(2001).
 RL EMBL, AK019981; BAB31950.1; -
 DR MGD; MGI:1924832; 5730589L02RIK.
 DR InterPro: IPR004299; MBOAT_fam.
 DR Pfam: PF03062; MBOAT_1.
 SQ SEQUENCE 473 AA; 53504 MW; CEGF8E93C3D01C4F CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;
 Best Local Similarity 100.0%; Pred. No. 42;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACGL 72
 |||||||
 DB 436 LALACGL 443

RESULT 29

ID O8R1P9 PRELIMINARY: PRT: 473 AA.
 AC O8R1P9:
 DT 01-JUN-2002 (TREMBLrel. 21, Created)
 DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE RIKEN CDNA 5730589L02 gene.
 OS Mus musculus (Mouse).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCB1_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 DR Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.

DR EMBL; BC023417; AAH23417.1; -
 SQ SEQUENCE 473 AA; 53382 MW; DAA1FE0DA78013EA CRC64;

Query Match 2.8%; Score 8; DB 11; Length 473;
 Best Local Similarity 100.0%; Pred. No. 42;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 65 LALACGL 72
 |||||||
 DB 436 LALACGL 443

RESULT 30

ID O9L1W0 PRELIMINARY: PRT: 522 AA.
 AC O9L1W0:
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
 DE Similar to an Arabidopsis thaliana chromosome BAC genomic
 DE sequence.
 OS Oryza sativa (Rice).
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 CC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 CC Euphorbiaceae; Oryzae; Oryza.
 OX NCB1_TaxID=4530;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hsing Y.C., Chow T., Chen C., Wu H., Chu M., Chao Y., Liu S.;
 RT "Oryza sativa PAC P0699E04 genomics sequence, complete sequence."
 RL Submitted (Jan-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP001111; BAA90509.1; -
 SQ SEQUENCE 522 AA; 54697 MW; 21C6BAD2441B56BF CRC64;

Query Match 2.8%; Score 8; DB 10; Length 522;
 Best Local Similarity 100.0%; Pred. No. 46;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRGE 50
 |||||||
 DB 415 RRRGRGE 422

RESULT 31

ID O9YTU9 PRELIMINARY: PRT: 564 AA.
 AC O9YTU9:
 DT 01-MAY-1999 (TREMBLrel. 10, Created)
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hemagglutinin H5 (Fragment).
 OS Influenzavirus A.
 CC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
 CC Influenza A viruses.
 OX NCB1_TaxID=11320;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A/DUCK/PORTSDAM/2216-4/84;
 RX MEDLINE=99099002; PubMed=9882316;
 RA Matrosovich M., Zhou N., Kawaoka Y., Webster R.;
 RT "The surface glycoproteins of H5 influenza viruses isolated from
 RT humans, chickens, and wild aquatic birds have distinguishable
 RT properties."
 RL J. Virol. 73:1146-1155(1999).
 CC -1- FUNCTION: HEMAGGLUTININ IS RESPONSIBLE FOR ATTACHING THE VIRUS TO
 CC CELL RECEPTORS AND FOR INITIATING INFECTION (BY SIMILARITY).
 CC -1- SUBUNIT: HOMOTRIMER. EACH OF THE MONOMERS IS FORMED BY TWO CHAINS
 CC (HA1 AND HA2) LINKED BY A DISULFIDE BOND (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE INFLUENZA HEMAGGLUTININ FAMILY.
 DR EMBL; AF082041; AAD13573.1; -
 DR HSSP; P03437; 1HTM.
 DR InterPro: IPR001364; Hemagglutn.
 DR Pfam: PF00509; Hemagglutinin; 1.

DR PRINTS: PR00329; HEMAGGLUTININ2.
DR PRODOM: PD000225; Hemagglutin; 1.
KW Envelope protein; Glycoprotein; Hemagglutinin.
FT NON_TER 564
SQ SEQUENCE 564 AA; 63562 MW; B317179A7F3E6F98 CRC64;

Query Match 2.8%; Score 8; DB 12; Length 564;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVSL 79
Db 6 LLLAVVSL 13

RESULT 32

OSOPLO PRELIMINARY; PRT; 568 AA.
AC OSOPLO;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hemagglutinin (Fragment).
GN HA.
OS Influenza A virus (A/Goose/Hong Kong/3014.8/2000(H5N1)).
OC Viruses; ssRNA negative-strand viruses; Orthomyxoviridae;
OC Influenza A viruses; Influenzavirus A.
OX NCBI_TaxID=176675;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=A/GOOSE/HONG KONG/3014.8/2000(H5N1);
RX MEDLINE=21874832; PubMed=11878904;
RA Guan Y., Peiris M., Kong K.F., Dyrting K.C., Ellis T.M., Sit T.,
Zhang L.J., Shortridge K.F.;
RT "H5N1 Influenza Viruses Isolated from Geese in Southeastern China:
Evidence for Genetic Reassortment and Interspecies Transmission to
Ducks.";
RT Virology 292:16-23(2002).
DR EMBL: AY059482; AL31388.1; .
FT NON_TER 568
SQ SEQUENCE 568 AA; 64281 MW; OB0A4CFE034F1769 CRC64;

Query Match 2.8%; Score 8; DB 12; Length 568;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 72 LLLAVVSL 79
Db 6 LLLAVVSL 13

RESULT 33

OSOPLO PRELIMINARY; PRT; 656 AA.
AC OSOPLO;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-OCT-2000 (TREMBlrel. 15, Last annotation update)
DE Possible hypothetical 61.9 kDa protein.
GN CHRI.338.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TREU927;
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
Gerrard C., Rajandream M.A., Barrell B.G.;
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL359782; CAB95571.1; .
FT NON_TER 656
SQ SEQUENCE 656 AA; 72138 MW; CBAC892D25937FAD CRC64;

Query Match 2.8%; Score 8; DB 5; Length 656;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 93 BELVAEED 100
Db 454 BELVAEED 461

RESULT 34

OSOPLO PRELIMINARY; PRT; 1523 AA.
AC OSOPLO;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Possibly hypothetical protein 85.6 kD.
GN CHRI.139.
OS Trypanosoma brucei.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
OX NCBI_TaxID=5691;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TREU927;
RA Hall N., Bowman S., Quail M., Ivens A.C., Kay M.P., Bray-Allen S.,
Lennard N.J., Clark L.N., Harris B.R., Melville S., Lawson D.,
Gerrard C., Rajandream M.A., Barrell B.G.;
RT Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL359782; CAB95435.1; .
KW Hypothetical protein.
SQ SEQUENCE 1523 AA; 168322 MW; 077BDC751CD1E5A CRC64;

Query Match 2.8%; Score 8; DB 5; Length 1523;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 218 VDGVALRL 225
Db 382 VDGVALRL 389

RESULT 35

OSOPLO PRELIMINARY; PRT; 24 AA.
AC OSOPLO;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Sperm protamine P3 (P03) (Fragment).
OS Octopus vulgaris (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC Incirrata; Octopodidae; Octopus.
OX NCBI_TaxID=6645;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=SPERM;
RA Gimenez-Bonate P., Ribes E., Buena C., Sautiere P., Kouach M.,
Aurio J., Kasinsky H.E., Chiva M.;
RT "Chromatin remodeling and protamines during spermiogenesis of Octopus
vulgaris (Cephalopoda).";
RT J. Exp. Zool. 0:0-0(2001).
DE -1- FUNCTION: PROTEIN SUBSTRATE FOR HISTONES IN THE CHROMATIN OF
SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- MASS SPECTROMETRY: MW=4389; METHOD=ELECTROSPRAY.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
DNA condensation; Nuclear protein.
FT DOMAIN 1 16
FT NON_TER 24
SQ SEQUENCE 24 AA; 3381 MW; 308E90ED9D2C9C9C CRC64;

```

Query Match          2.5%; Score 7; DB 5; Length 24;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49
    |||||
Db 11 RRRGRG 17

RESULT 36
P83217 PRELIMINARY; PRT: 30 AA.
AC P83217:
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DE 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Sperm protamine p5 (Po5).
OS Octopus vulgaris (Octopus).
OC Eukaryota; Metazoa; Mollusca; Cephalopoda; Coleoidea; Octopoda;
OC Incurrata; Octopodidae; Octopus.
OX NCBI_TaxID=6645;
RN [1]
RP SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.
RC TISSUE=SPERM;
RA Gimenez-Bonafe P., Ribes E., Buesa C., Sautiere P., Konach M.,
RA Ausio J., Kasinsky H.E., Chiva M.;
RT "Chromatin remodeling and protamines during spermiogenesis of Octopus
RT vulgaris (Cephalopoda).";
RL J. Exp. Zool. 0:0-0(2001).
CC -1- FUNCTION: PROTAGINIS SUBSTITUTE FOR HISTONES IN THE CHROMATIN OF
CC SPERM DURING THE HAPLOID PHASE OF SPERMATOGENESIS. THEY COMPACT
CC SPERM DNA INTO A HIGHLY CONDENSED, STABLE AND INACTIVE COMPLEX.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: TESTIS.
CC -1- MASS SPECTROMETRY: MW=3941; METHOD=ELECTROSPRAY.
KW Chromosomal protein; Nucleosome core; Spermatogenesis; DNA-binding;
KW DNA condensation; Nuclear protein.
FT DOMAIN 2 15 POLY-ARG.
FT DOMAIN 17 26 POLY-ARG.
SQ SEQUENCE 30 AA; 3943 MW; 14F1BC7E4D277049 CRC64;

Query Match          2.5%; Score 7; DB 5; Length 30;
Best Local Similarity 100.0%; Pred. No. 35;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 43 RRRGRG 49
    |||||
Db 5 RRRGRG 11

RESULT 37
Q9BEF6 PRELIMINARY; PRT: 32 AA.
ID Q9BEF6:
AC Q9BEF6:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Beta-lactoglobulin precursor (Fragment).
GN BETA-LACTOGLOBULIN.
OS Capra hircus (Goat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Capra.
OX NCBI_TaxID=9925;
RN [1]
RP SEQUENCE FROM N.A.
RA D'Andrea M., Pilla F., Graziano M.;
RT "A new polymorphism of goat beta-lactoglobulin proximal promoter
RT region.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AJ292058; CAC27455.1; -.
KW Signal.
FT SIGNAL 1 18 POTENTIAL.

```

```

FT NON-TER 32 32
SQ SEQUENCE 32 AA; 3372 MW; 0C56BD579B30C190 CRC64;

Query Match          2.5%; Score 7; DB 6; Length 32;
Best Local Similarity 100.0%; Pred. No. 37;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 63 LGLALAC 69
    |||||
Db 8 LGLALAC 14

RESULT 38
Q9ZG35 PRELIMINARY; PRT: 35 AA.
ID Q9ZG35:
AC Q9ZG35:
DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DE 01-MAY-1999 (TREMBLrel. 10, Last annotation update)
DE Hypothetical 3.5 kDa protein (Fragment).
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=L2 434B;
RA Wang L., Steenburg S.D., Zheng Y., Larsen S.H.;
RT "Gene identification of Chlamydia trachomatis by random DNA
RT sequencing.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF087333; AAD04107.1; -.
KW Hypothetical protein.
FT NON-TER 1 35
FT NON-TER 35 35
SQ SEQUENCE 35 AA; 3519 MW; 4668B72BAF28F7D CRC64;

Query Match          2.5%; Score 7; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 NSSSPLR 180
    |||||
Db 8 NSSSPLR 14

RESULT 39
O16193 PRELIMINARY; PRT: 54 AA.
ID O16193:
AC O16193:
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Orf2 5' to PD-ECGF/TP protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94312438; PubMed=8038210;
RA Usuki K., Gomez L.J., Wernstedt C., Moren A., Miyazono K.,
RA Claesson-Welsh L., Heldin C.H.;
RT "Structural properties of 3.0 kb and 3.2 kb transcripts encoding
RT platelet-derived endothelial cell growth factor/thymidine
RT phosphorylase in A431 cells.";
RL Biochim. Biophys. Acta 1222:411-414(1994).
DR EMBL; S72487; AAD14107.1; -.
SQ SEQUENCE 54 AA; 5808 MW; 827925FBA70CB222 CRC64;

Query Match          2.5%; Score 7; DB 4; Length 54;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 125 RPRRSAP 131
 |||||
 Db 17 RPRRSAP 23

RESULT 40

Q8R0S4 PRELIMINARY; PRT; 67 AA.
 ID Q8R0S4
 AC Q8R0S4;
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DE Hypothetical 7.8 kDa protein.
 GN YHCR.
 OS Serratia marcescens.
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 OC Serratia.
 OX NCBI_TaxID=615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Baba M., Midorikawa Y., Nakagawa Y., Fujita M., Matsuyama T.;
 RT "Serratia marcescens and Escherichia coli genes controlling
 temperature-dependent production of structurally unrelated secondary
 metabolites such as prodigiosin and serrawettin.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB080601; BAB85653.1; -.
 KM Hypothetical protein
 SQ SEQUENCE 67 AA: 7791 MM; 47B01A87E69AC2A2 CRC64;

Query Match 2.5%; Score 7; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 72;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 55 LVPPLAL 61
 |||||
 Db 21 LVPPLAL 27

Search completed: March 31, 2003, 14:15:52
 Job time : 88 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 05:28:45 ; Search time 53.5639 Seconds

(without alignments)
706.506 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDPEISARRLPRLPSLUG.....PNAHLKAPFLRYFGLFQVH 284

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

A.Geneseq-101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1444	100.0	284	19	AAW47525
2	1268	87.8	249	20	AAW09369
3	1268	87.8	249	21	AAW07526
4	1268	87.8	249	21	AAW5338
5	1268	87.8	249	23	AAW86129
6	1265	87.6	249	19	AAW29745
7	1265	87.6	249	22	AAW00891
8	1066	73.8	273	22	AAW03499
9	1062	73.5	208	20	AAW93590
10	1020	70.6	225	19	AAW47524

11	1020	70.6	225	21	AAW07527
12	968	67.0	211	20	AAW93591
13	792	54.8	189	19	AAW29746
14	792	54.8	189	22	AAE00892
15	761	52.7	146	22	AAE00895
16	116	8.0	325	22	AAE67553
17	116	8.0	409	23	AAU77718
18	108.5	7.5	211	21	AAW58216
19	108.5	7.5	260	21	AAW58215
20	106.5	7.4	254	16	AAW64190
21	106.5	7.4	254	18	AAW26657
22	106.5	7.4	254	23	AAW5953
23	106	7.3	1428	21	AAW97033
24	105.5	7.3	406	23	AAU77717
25	104.5	7.2	779	23	AAW07845
26	104	7.2	409	23	AAU77716
27	104	7.2	1323	15	AAW5248
28	102	7.1	256	22	AAW5248
29	100.5	7.0	647	17	AAW04327
30	100	6.9	220	22	AAW62340
31	99	6.9	574	21	AAW97032
32	99	6.9	1008	22	AAW78891
33	99	6.9	1020	22	AAW79875
34	96.5	6.7	1097	22	ABG25655
35	96.5	6.7	1631	22	ABG22481
36	96	6.6	234	22	AAW62339
37	95	6.6	240	23	AAE13680
38	95	6.6	294	19	AAW6956
39	95	6.6	294	19	AAW68292
40	95	6.6	294	22	AAE08737
41	95	6.6	294	22	AAE04425
42	95	6.6	294	22	AAE01992
43	94.5	6.5	1006	22	ABG21178
44	94	6.5	876	22	ABG00217
45	93.5	6.5	409	22	AAW00076

ALIGNMENTS

RESULT 1	AAW47525	AAW47525 standard; Protein: 284 AA.
ID	AAW47525	
AC	AAW47525	
XX		
DT	21-JUL-1998	(first entry)
XX		
DE	Homo sapiens tumour necrosis factor related ligand (TRELL).	
XX		
KW	TRELL, tumour necrosis factor related ligand; tnfr; treatment; cancer; autoimmune disease; immune system; stimulation; suppression; graft rejection.	
KW		
XX		
OS	Homo sapiens.	
XX		
PN	WO9805783-A1.	
PD	12-FEB-1998.	
XX		
PF	07-AUG-1997;	97WO-US13945.
XX		
PR	18-MAR-1997;	97US-0040820.
PR	07-AUG-1996;	96US-0023541.
XX	18-OCT-1996;	96US-0028515.
XX		
PA	(BIOJ) BIOGEN INC.	
XX	(UYGE-) UNIV GENEVA FACULTY MEDICINE.	
XX		
PI	Browning JL, Chicopeportliche Y;	
XX		
DR	WPI, 1998-145619/13.	
DR	N-PSDB: AAV18600.	

Amino acid sequenc
Mouse TNRL3 protei
TNF related endoth
Human UFA1fag TREP
Human TREPA (TNF r
Drosophila melanog
Drosophila melanog
Canine mature CD15
Canine CD154, Can
Human 4-1BB-L poly
Human 4-1BB ligand
Human cytokine 4-1
Caspase 8-interact
Drosophila melanog
Human MAP kinase p
Drosophila melanog
N-methyl-D-aspart
Rat petrin. Rattu
Gp120 V3 loop-CD15
Caspase 8-interact
Human protein SEQ
Human protein SEQ
Novel human diagno
Novel human diagno
Gp120 V3 loop-CD15
Human HIVM-binding
NF-kB receptor act
NF-kB receptor act
Murine receptor ac
Murine receptor ac
Murine RANKL (rece
Novel human diagno
Novel human diagno
Human polypeptide

```
XX Tumour necrosis factor related ligand - useful for, e.g. treating
PT cancer, auto-immune disease and immune responses to tissue grafts
XX Claim 12; Pages 50-51; 69pp; English.
PS
XX
XX The sequence is that of human tumour necrosis factor related
CC ligand (TRELL). TRELL or active fragments can be included with a
CC carrier in pharmaceutical compositions to treat cancer, autoimmune
CC diseases or immune responses to tissue grafts, or to stimulate or
CC suppress the immune system. It is useful to screen for TRELL
CC receptors, by labelling with a detectable label and screening
CC compositions for binding. Agents interfering with TRELL-receptor
CC binding can also be screened for, can then be administered,
CC optionally with interferon- gamma, to induce cell death or
CC treat, suppress or alter immune responses (especially involving human
CC adenocarcinoma cells) involving a signal pathway between TRELL and its
CC receptor. It's coding sequence can be used in gene therapy for
CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
CC autoimmune and inflammatory diseases or inherited genetic disorders,
CC by introducing into cells, e.g. a virus comprising a gene encoding TRELL.
CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
CC It may also be of use in the preparation of prepare probes for
CC screening natural/synthetic DNAs for TRELL-encoding sequences
CC and for antisense therapy.
CC
CC
XX
SQ Sequence 284 AA:
Query Match 100.0%; Score 1444; DB 19; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.2e-129;
Matches 284; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 MSLLPEISARLPLPRSGISRDGAVRQAPPAARRRSGRRGEPGTALLVPLA 60
DB 1 MSLLPEISARLPLPRSGISRDGAVRQAPPAARRRSGRRGEPGTALLVPLA 60
OY 61 LGLGLACGLGLLAVVSLGSRASLSAOPAEELVAEEDDPSSELNPTDESSODAPPL 120
DB 61 LGLGLACGLGLLAVVSLGSRASLSAOPAEELVAEEDDPSSELNPTDESSODAPPL 120
OY 121 NRLVPRRSAPKGRKTRARRAIAAHYEVHPRPGODGAQAGVGTVSGWEARINSSPLR 180
DB 121 NRLVPRRSAPKGRKTRARRAIAAHYEVHPRPGODGAQAGVGTVSGWEARINSSPLR 180
OY 181 YNRQIGEFIVTRAGLYLYLCQVHFDEGKAVYLKLDLVGVALRCLLEFSATTAASLGP 240
DB 181 YNRQIGEFIVTRAGLYLYLCQVHFDEGKAVYLKLDLVGVALRCLLEFSATTAASLGP 240
OY 241 QLRICOVSGLLALRPGSSLRIRTLPMANHLKAAPFLTYEGLFQVH 284
DB 241 QLRICOVSGLLALRPGSSLRIRTLPMANHLKAAPFLTYEGLFQVH 284
RESULT 2
AAY09369
ID AAY09369 standard; Protein: 249 AA.
XX
XX AAY09369;
XX
XX 15-JUL-1999 (first entry)
XX
XX Human tumour necrosis factor Apo-3 ligand protein sequence.
XX
XX Human; tumour necrosis factor; Apo-3 ligand; lymphotoxin; apoptosis;
XX NF-kappaB-dependent transcription; JNK/SAPK-dependent response;
XX cancer.
XX
XX Homo sapiens.
XX
XX OS
XX PN MO9919490-A1.
XX PD 22-APR-1999.
XX
```

```
PF 09-OCT-1998; 98WO-US21407.
XX
XX 17-DEC-1997; 97US-0069862.
XX 10-OCT-1997; 97US-0062037.
XX
XX
XX (GENTH) GENENTECH INC.
XX
XX Ashkenazi AJ, Marsters SA, Pitter R;
XX WPI; 1999-287982/24.
XX N-PSDB; AAX56000.
XX
XX New human Apo-3 ligand (a tumour necrosis factor) homologue
XX
XX Claim 1; Fig 1; 74pp; English.
XX
XX The present sequence represents a human tumour necrosis factor (TNF)
XX and lymphotoxin homologue designated Apo-3 ligand. Apo-3 ligand has
XX cytostatic activity. Apo-3 ligand can be used to induce apoptosis in
XX mammalian cancer cells, to induce NF-kappaB-dependent transcription and
XX to induce JNK/SAPK-dependent responses in mammalian cells.
XX
XX
XX
SQ Sequence 249 AA:
Query Match 87.8%; Score 1268; DB 20; Length 249;
Best Local Similarity 100.0%; Pred. No. 6.6e-113;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 36 MAARRSORRRGRGEPGTALLVPLALGGLALACGLLAVVSLGSRASLSAOPAEEL 95
DB 1 MAARRSORRRGRGEPGTALLVPLALGGLALACGLLAVVSLGSRASLSAOPAEEL 95
OY 96 VAEEDDPSSELNPTDESSODAPPLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGOD 155
DB 96 VAEEDDPSSELNPTDESSODAPPLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGOD 155
OY 61 VAEEDDPSSELNPTDESSODAPPLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGOD 120
DB 61 VAEEDDPSSELNPTDESSODAPPLNRLVPRRSAPKGRKTRARRAIAAHYEVHPRGOD 120
OY 156 GAQAGVDGTVSGWEARINSSPLRNRQIGEFIVTRAGLYLYLCQVHFDEGKAVYLKLD 215
DB 156 GAQAGVDGTVSGWEARINSSPLRNRQIGEFIVTRAGLYLYLCQVHFDEGKAVYLKLD 215
OY 121 GAQAGVDGTVSGWEARINSSPLRNRQIGEFIVTRAGLYLYLCQVHFDEGKAVYLKLD 180
DB 121 GAQAGVDGTVSGWEARINSSPLRNRQIGEFIVTRAGLYLYLCQVHFDEGKAVYLKLD 180
OY 216 LLYDGVVALRCLLEFSATTAASSLGQRLRCOVSGLLALRPGSSLRIRTLPMANHLKAAPL 275
DB 181 LLYDGVVALRCLLEFSATTAASSLGQRLRCOVSGLLALRPGSSLRIRTLPMANHLKAAPL 240
OY 276 TYEGLFQVH 284
DB 241 TYEGLFQVH 249
RESULT 3
AAB07526
ID AAB07526 standard; Protein: 249 AA.
XX
XX AAB07526;
XX
XX 20-OCT-2000 (first entry)
XX
XX Amino acid sequence of a soluble recombinant human TWEAK protein.
XX
XX TWEAK protein; immunological disorder; immune response; inflammation;
XX TWEAK blocking agent; autoimmune disease; organ transplant rejection;
XX Graft-versus-Host disease; GVHD; lymphoid cell malignancy; shock; tumour.
XX
XX Homo sapiens.
XX
XX OS
XX PN WO200042073-A1.
XX PD 20-JUL-2000.
XX
XX 14-JAN-2000; 2000WO-US01044.
XX PR 15-JAN-1999; 99US-0116168.
XX
```


Best Local Similarity 100.0%; Pred. No. 6.6e-113;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 36 MAARSRRRGRGREGPTALVPLALGIGLALACGLLAVVSLGSRASLSAQEPAGEEL 95
DB 1 MAARSRRRGRGREGPTALVPLALGIGLALACGLLAVVSLGSRASLSAQEPAGEEL 60
OY 96 VAEEDODPSLELNPQTEESODPAFLNRLVPRRSAPKGGKKTARRAIAAHYEVHPPGCD 155
DB 61 VAEEDODPSLELNPQTEESODPAFLNRLVPRRSAPKGGKKTARRAIAAHYEVHPPGCD 120
OY 156 GAQAGVDGTVSGMEERARINSSPLRNQIGEFYTRAGLYLYLCOVHDEGKAVYLKLD 215
DB 121 GAQAGVDGTVSGMEERARINSSPLRNQIGEFYTRAGLYLYLCOVHDEGKAVYLKLD 180
OY 216 LLVDGVALARCLCEFSATTAASSLGPQLRCQVSGLLALRPSSLRIRTLPMANHLKAAPFL 275
DB 181 LLVDGVALARCLCEFSATTAASSLGPQLRCQVSGLLALRPSSLRIRTLPMANHLKAAPFL 240
OY 276 TYFGLFQVH 284
DB 241 TYFGLFQVH 249

```

RESULT 5

AAU86129
ID AAU86129 standard; Protein: 249 AA.

```

AC AAU86129;
DT 15-JUL-2002 (first entry)
DE Human PRO207 polypeptide.
XX
KW Human; PRO: benign tumour; malignant tumour; lymphoid malignancy;
KW leukaemia; neuronal disorder; stromal disorder; blastocoele disorder;
KW inflammatory disorder; immune disorder; angiogenic disorder;
KW cytoskeletal; neuroprotective.
XX
OS Homo sapiens.
XX
PN WO200153486-A1.
PD 26-JUL-2001.
XX
PF 11-FEB-2000; 2000WO-US03565.
XX
PR 08-MAR-1999; 99WO-US05028.
PR 11-MAR-1999; 99US-123972P.
PR 11-MAY-1999; 99US-133459P.
PR 02-JUN-1999; 99WO-US12252.
PR 22-JUN-1999; 99US-140650P.
PR 22-JUN-1999; 99US-140653P.
PR 20-JUL-1999; 99US-144758P.
PR 26-JUL-1999; 99US-145698P.
PR 28-JUL-1999; 99US-146222P.
PR 17-AUG-1999; 99US-149395P.
PR 31-AUG-1999; 99US-151689P.
PR 01-SEP-1999; 99WO-US20111.
PR 15-SEP-1999; 99WO-US21090.
PR 30-NOV-1999; 99WO-US28313.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 05-JAN-2000; 2000WO-US00219.
XX
PA (GETH ) GENENTECH INC.
XX
PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Hillan KJ;
PI Marsters SA, Pan J, Pitter RM, Roy MA, Smith V, Stone DM;
PI Watanabe CK, Wood WI;
XX
DR WPI: 2002-205567/26.
DR N-PSDB: ABK40255.
XX

```

XX Thirty five nucleic acids encoding PRO polypeptides, useful for
PT treating benign or malignant tumours, leukaemias and lymphoid
PT malignancies, inflammatory, angiogenic and immunologic disorders -
XX
PS Claim 61; Fig 4; 302pp; English.

XX The present invention relates to the isolation of novel human PRO
CC polypeptides and the polynucleotide sequences encoding them. The
CC PRO polypeptides, agonists, antagonists or anti-PRO antibodies are
CC useful for treating benign or malignant tumours (e.g. renal, kidney,
CC bladder, breast, etc), leukaemias and lymphoid malignancies, other
CC disorders such as neuronal, glial, astrocytal, hypothalamic, glandular,
CC macrophagal, stromal and blastocoele disorders, inflammatory, immune
CC and angiogenic disorders. The polynucleotide sequences are also
CC useful in gene therapy. AAU86128-AAU86162 represent the human PRO
CC polypeptides of the invention.

XX Sequence 249 AA:

Query Match 87.8%; Score 1268; DB 23; Length 249;

Best Local Similarity 100.0%; Pred. No. 6.6e-113;
Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 36 MAARSRRRGRGREGPTALVPLALGIGLALACGLLAVVSLGSRASLSAQEPAGEEL 95
DB 1 MAARSRRRGRGREGPTALVPLALGIGLALACGLLAVVSLGSRASLSAQEPAGEEL 60
OY 96 VAEEDODPSLELNPQTEESODPAFLNRLVPRRSAPKGGKKTARRAIAAHYEVHPPGCD 155
DB 61 VAEEDODPSLELNPQTEESODPAFLNRLVPRRSAPKGGKKTARRAIAAHYEVHPPGCD 120
OY 156 GAQAGVDGTVSGMEERARINSSPLRNQIGEFYTRAGLYLYLCOVHDEGKAVYLKLD 215
DB 121 GAQAGVDGTVSGMEERARINSSPLRNQIGEFYTRAGLYLYLCOVHDEGKAVYLKLD 180
OY 216 LLVDGVALARCLCEFSATTAASSLGPQLRCQVSGLLALRPSSLRIRTLPMANHLKAAPFL 275
DB 181 LLVDGVALARCLCEFSATTAASSLGPQLRCQVSGLLALRPSSLRIRTLPMANHLKAAPFL 240
OY 276 TYFGLFQVH 284
DB 241 TYFGLFQVH 249

```

RESULT 6

AAW29745
ID AAW29745 standard; Protein: 249 AA.

```

AC AAW29745;
DT 27-OCT-1998 (first entry)
DE TNF related endothelium proliferative agent protein.
XX
KW TNF; endothelium proliferative agent; TRBP; wound healing; cancer;
KW tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
XX
OS Homo sapiens.
XX
PN WO9835061-A2.
PD 13-AUG-1998.
PF 12-FEB-1998; 98WO-US02859.
XX
PR 10-FEB-1998; 98US-0021706.
PR 12-FEB-1997; 97US-0798692.
XX
PA (ABBO ) ABBOTT LAB.
XX
PI Willey SR;
XX

```

DR WPI: 1998-447255/38.
 DR N-PSDB; AAV47613.
 XX
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
 PT treatment of autoimmune disease, tumours and inflammation
 XX
 PS Claim 16; Page 123-4; 142pp; English.
 XX
 CC The TNF-related endothelium proliferative agent (TREPA), or its
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
 CC promote wound healing or tissue grafting, by promoting vascularisation,
 CC also to induce apoptosis for treating cancer and eliminating autoreactive
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
 CC TREPA peptides can also be used to target cytotoxic agents or for
 CC affinity isolation of the corresponding receptor, the nucleic acid for
 CC which can be used to transform tumour cells to render them more
 CC responsive to TREPA and to screen for TREPA mimics.
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
 CC vascularisation), inflammation or a wide range of autoimmune conditions,
 CC conditions involving abnormal stimulation of epithelial cells (e.g.
 CC atherosclerosis), for birth control (inhibiting ovulation and placental
 CC formation) or other angiogenic conditions (e.g. ulcers).
 CC
 XX
 SQ Sequence 249 AA:
 Query Match 87.6%; Score 1265; DB 19; Length 249;
 Best Local Similarity 99.6%; Pred. No. 1.3e-112;
 Matches 248; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 36 MAARRSORRRGRGEPGTALLVPLALGLALACLGILLAVVSGSRASISAQEPAGEEL 95
 DB 1 MAARRSORRRGRGEPGTALLVPLALGLALACLGILLAVVSGSRASISAQEPAGEEL 60
 QY 96 VAEEDQPSSELNPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOD 155
 DB 61 VAEEDQPSSELNPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOD 120
 QY 156 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFVTFRAGLYLYLCQVHFDGKAVYTKLD 215
 DB 121 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFVTFRAGLYLYLCQVHFDGKAVYTKLD 180
 QY 216 LTVDGVALACLSEFSATASSISGLPQLRICOVSGLLALRPSSLRITLPPAHILKAAPFL 275
 DB 181 LTVDGVALACLSEFSATASSISGLPQLRICOVSGLLALRPSSLRITLPPAHILKAAPFL 240
 QY 276 TYFGLEFOVH 284
 DB 241 TYFGLEFOVH 249
 RESULT 7
 AAE00891
 ID AAE00891 standard; Protein; 249 AA.
 XX
 AC AAE00891;
 XX
 DT 04-JUL-2001 (first entry)
 XX
 DE Human TREPA (TNF related endothelium proliferative agent).
 XX
 KW Human; tumour necrosis factor; TNF; angiogenesis; wound healing;
 KW TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
 KW grafting; vulnerability.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 98..249
 FT /label= Extracellular_domain
 XX
 PN US6207642-B1.
 XX

PD 27-MAR-2001.
 XX
 PF 26-JUN-1998; 98US-0105343.
 XX
 PR 12-FEB-1997; 97US-0798692.
 PR 10-FEB-1998; 98US-0021706.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PT Wiley SR;
 XX
 DR WPI: 2001-280760/29.
 DR N-PSDB; AAD04350.
 XX
 PT Inducing angiogenesis in mammal at desired sites for promoting wound
 PT healing, by administering soluble fragment of extracellular domain of
 PT tumor necrosis factor related endothelium proliferative agent protein
 XX
 PS Claim 1; Column 75-76; 53pp; English.
 XX
 CC The present invention relates to extracellular signal molecules,
 CC particularly members of tumour necrosis factor (TNF) family molecules
 CC designated as TREPA (TNF related endothelium proliferative agent).
 CC Soluble biologically active TREPA are used to treat TREPA-associated
 CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
 CC in human for promoting wound healing and for vascularising grafted
 CC tissue for successful grafting and to promote tissue grafts.
 CC The present amino acid sequence is clone ID #690050 human TREPA.
 CC
 XX
 SQ Sequence 249 AA:
 Query Match 87.6%; Score 1265; DB 22; Length 249;
 Best Local Similarity 99.6%; Pred. No. 1.3e-112;
 Matches 248; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 QY 36 MAARRSORRRGRGEPGTALLVPLALGLALACLGILLAVVSGSRASISAQEPAGEEL 95
 DB 1 MAARRSORRRGRGEPGTALLVPLALGLALACLGILLAVVSGSRASISAQEPAGEEL 60
 QY 96 VAEEDQPSSELNPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOD 155
 DB 61 VAEEDQPSSELNPOTESODPAPFLNLRVPRRSAPKGRKTRARRAIAHYEVHPRGOD 120
 QY 156 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFVTFRAGLYLYLCQVHFDGKAVYTKLD 215
 DB 121 GAQAGVDGTVSGWEARINSSPLRYNRQIGEFVTFRAGLYLYLCQVHFDGKAVYTKLD 180
 QY 216 LTVDGVALACLSEFSATASSISGLPQLRICOVSGLLALRPSSLRITLPPAHILKAAPFL 275
 DB 181 LTVDGVALACLSEFSATASSISGLPQLRICOVSGLLALRPSSLRITLPPAHILKAAPFL 240
 QY 276 TYFGLEFOVH 284
 DB 241 TYFGLEFOVH 249
 RESULT 8
 AAU03499
 ID AAU03499 standard; Protein; 273 AA.
 XX
 AC AAU03499;
 XX
 DT 26-SEP-2001 (first entry)
 XX
 DE TWEAK extracellular domain-containing fusion protein.
 XX
 KW TWEAK extracellular domain; tumour necrosis factor; TNF; angiogenesis;
 KW ocular neovascularisation; diabetic retinopathy; neovascular glaucoma;
 KW retinoblastoma; retinopathy of prematurity; retrolental fibroplasia;
 KW rubeosis; uveitis; macular degeneration; arthritis; rheumatism;
 KW corneal graft neovascularisation; psoriasis; metastatic condition;
 KW malignant tumour; sarcoma; carcinoma; benign tumour; haemophilic joint;
 KW

KM preneoplastic condition; myocardial angiogenesis; wound granulation;
 KM scleroderma; vascular adhesion; telangiectasia; ischemia; human;
 KM atherosclerotic plaque neovascularisation; coronary atherosclerosis;
 KM peripheral atherosclerosis; pDC409-L2-TWEAK; TWEAK receptor; TWEAK;
 fusion protein.
 XX Homo sapiens.
 OS Synthetic.
 XX MO200145730-A2.
 XX 28-JUN-2001.
 XX 19-DEC-2000; 2000WO-US34755.
 XX 20-DEC-1999; 99US-0172878.
 PR 10-MAY-2000; 2000US-0203347.
 XX (IMMV) IMMUNEX CORP.
 XX Willey SR;
 PI WPI: 2001-417975/44.
 DR N-PSDB; AAS03964.
 XX Modulating angiogenesis in a mammal for treating diseases mediated by
 PT angiogenesis, e.g. solid tumours and vascular deficiencies of cardiac or
 PT peripheral tissue, by administering antagonist or agonist of TWEAK
 PT receptor
 XX Example 1; Page 41; 46pp; English.
 PS The sequence represents a fusion protein encoded by the
 XX expression vector pDC409-L2-TWEAK. The fusion protein comprises a
 CC growth hormone leader, a leucine zipper multimerisation domain, and
 CC the extracellular domain of human TWEAK. The fusion protein was
 CC used in the isolation of human TWEAK receptor (TWEAKR)-expressing
 CC clones from a COS cell human cDNA library. The TWEAK protein is
 CC a member of the tumor necrosis factor (TNF) family and induces
 CC angiogenesis. TWEAKR may therefore be used to screen for and
 CC develop TWEAKR agonists and antagonists for the modulation of
 CC angiogenesis, to be used in the treatment and diagnosis of human disease.
 CC The disorders mediated by angiogenesis include ocular disorders
 CC characterised by ocular neovascularisation such as diabetic retinopathy,
 CC neovascular glaucoma, retinoblastoma, retinopathy of prematurity,
 CC retrolental fibroplasia, rubecosis, uveitis, macular degeneration and
 CC corneal graft neovascularisation, and inflammatory diseases such as
 CC arthritis, rheumatism and psoriasis. Other treatable diseases include
 CC malignant and metastatic conditions such as sarcomas and carcinomas,
 CC benign tumours and preneoplastic conditions, myocardial angiogenesis,
 CC haemophilic joints, scleroderma, vascular adhesions, atherosclerotic
 CC plaque neovascularisation, telangiectasia, wound granulation, coronary
 CC atherosclerosis, peripheral atherosclerosis and ischemia.
 XX Sequence 273 AA;
 SO Query Match 73.8%; Score 1066; DB 22; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.5e-93;
 Matches 207; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 247 SLRITLPMWHLKAPFLTYFGLFQVH 273
 ||||||||||||||||||||||||||||
 RESULT 9
 AAW93590
 ID AAW93590 standard; Protein: 208 AA.
 XX
 AC AAW93590;
 XX
 DT 18-JUN-1999 (first entry)
 XX
 DE Human TNFR3 protein.
 XX
 KM Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KM developmental abnormality; gestational abnormality; prostate cancer;
 KM APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KM apoptosis; human.
 XX
 OS Homo sapiens.
 XX
 PM WO9911791-A2.
 PD 11-MAR-1999.
 XX
 PF 04-SEP-1998; 98WO-US18393.
 XX
 PR 05-SEP-1997; 97US-0924634.
 XX
 PA (UNIV) UNIV WASHINGTON.
 XX
 PI Chaudhary PM;
 XX
 DR WPI: 1999-205191/17.
 DR N-PSDB; AAX23424.
 XX
 PR New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 XX
 PS Claim 40; Fig 13A; 156pp; English.
 XX
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active
 CC fragments and APO4 signal transducer molecules that specifically interact
 CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
 CC activity. The method is performed in vivo or in vitro. APO polypeptides
 CC are all useful as immunogens for preparing antibodies. APO4 is also
 CC useful for diagnosis/treatment of developmental or gestational
 CC abnormalities. APO8 was transfected to human breast carcinoma cell line
 CC MCF-7, and induced apoptosis.
 XX Sequence 208 AA;
 SO Query Match 73.5%; Score 1062; DB 20; Length 208;
 Best Local Similarity 99.5%; Pred. No. 2.6e-93;
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

DB 1 VSIGSRASLSAEPQAEELVAEEDDPSSELNPQTEESQDPAPFLNRLVRRPRAKGRKT 60
 QY 137 RARRATAIAHYEVHPRPGODGAGVDSVSGWEARINSSPLRYNROIGEFYTRAGLY 196
 DB 61 RARRATAIAHYEVHPRPGODGAGVDSVSGWEARINSSPLRYNROIGEFYTRAGLY 120
 QY 197 YLYCOVHFDEGKAVYLYKLDLVGVYALRCLLEFSATPAASLGPQLRQVSGGLAARPG 256
 DB 121 YLYCOVHFDEGKAVYLYKLDLVGVYALRCLLEFSATPAASLGPQLRQVSGGLAARPG 180
 QY 257 SSIRIRFLPMHLLKAAPEFLTYFGFLQVH 284
 DB 181 SSIRIRFLPMHLLKAAPEFLTYFGFLQVH 208

RESULT 10
 ID AAM47524 standard; Protein; 225 AA.
 AC AAM47524;
 DT 21-JUL-1998 (first entry)
 DE Mus musculus tumour necrosis factor related ligand (TRELL).
 XX
 XX TRELL; tumour necrosis factor related ligand; tnf; treatment;
 KM cancer; autoimmune disease; immune system; stimulation; suppression;
 KM graft rejection.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 FT Domain 1..21
 FT /note= "hydrophobic, transmembrane domain"
 XX
 PN W09805783-A1.
 PD 12-FEB-1998.
 XX
 PF 07-AUG-1997; 97MO-US13945.
 XX
 PR 18-MAR-1997; 97US-0040820.
 PR 07-AUG-1996; 96US-0023541.
 PR 18-OCT-1996; 96US-0028515.
 XX
 XX
 PA (BIOJ) BIOGEN INC.
 PA (UYGE-) UNIV GENEVA FACULTY MEDICINE.
 XX
 PI Browning JL, Chicheportiche Y;
 XX
 DR WPI; 1998-145619/13.
 DR N-PSDB; AAV18599.
 XX
 PT Tumour necrosis factor related ligand - useful for, e.g. treating
 PT cancer, auto-immune disease and immune responses to tissue grafts
 XX
 XX
 PS Claim 12; Pages 48-50; 69pp; English.

CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
 CC It may also be of use in the preparation of prepare probes for
 CC screening natural/synthetic DNAs for TRELL-encoding sequences
 CC and for antisense therapy.
 CC
 XX
 XX Sequence 225 AA;
 SQ
 Query Match 70.6%; Score 1020; DB 19; Length 225;
 Best Local Similarity 88.8%; Pred. No. 2, 9e-89;
 Matches 199; Conservative 9; Mismatches 16; Indels 0; Gaps 0;
 QY 61 LGIGLALACGLLLAVVSLGSRASLSAEPQAEELVAEEDDPSSELNPQTEESQDPAPFL 120
 DB 2 LSLGLALACGLLLAVVSLGSMATLSAEPQAEELVAEEDDPSSELNPQTEESQDPAPFL 61
 QY 121 NRLVRRPRAKGRKTARRAIAHYEVHPRPGODGAGVDSVSGWEARINSSPLR 180
 DB 62 EQLVRRPRAKGRKTARRAIAHYEVHPRPGODGAGVDSVSGWEARINSSPLR 121
 QY 181 YNROIGEFYTRAGLYLYCOVHFDEGKAVYLYKLDLVGVYALRCLLEFSATPAASLGP 240
 DB 122 YDROIGEFYTRAGLYLYCOVHFDEGKAVYLYKLDLVGVYALRCLLEFSATPAASLGP 181
 QY 241 QLRLQVSGGLAARPGSSLRIRFLPMHLLKAAPEFLTYFGFLQVH 284
 DB 182 QLRLQVSGGLAARPGSSLRIRFLPMHLLKAAPEFLTYFGFLQVH 225

RESULT 11
 ID AAB07527 standard; Protein; 225 AA.
 AC AAB07527;
 DT 20-OCT-2000 (first entry)
 DE Amino acid sequence of a soluble recombinant murine TWEAK protein.
 XX
 XX TWEAK protein; Immunological disorder; immune response; inflammation;
 KM TWEAK blocking agent; autoimmune disease; organ transplant rejection;
 KM Graft-versus-host disease; GVHD; lymphoid cell malignancy; shock; tumour.
 XX
 OS Mus sp.
 XX
 PN W0200042073-A1.
 PD 20-JUL-2000.
 XX
 PF 14-JAN-2000; 2000MO-US01044.
 XX
 PR 15-JAN-1999; 99US-0116168.
 XX
 PA (BIOJ) BIOGEN INC.
 PA Renner P;
 XX
 PI Renner P;
 XX
 DR WPI; 2000-476036/41.
 XX
 PT Preventing and treating immune responses using modulators, especially
 PT antibodies, of TWEAK, TWEAK receptors and TWEAK ligands, useful for
 PT treating e.g. inflammation and graft versus host disease -
 XX
 PS Disclosure; Fig 1; 45pp; English.

CC The present sequence represents a TWEAK protein. The specification
 CC describes a method for preventing or treating an immunological
 CC disorder and/or inhibiting an immune response in an animal. The
 CC method comprises administering a TWEAK blocking agent. The method may
 CC be used for preventing and treating immune disorders associated with
 CC inappropriate expression and/or activity of TWEAK. These disorders
 CC include autoimmune diseases, acute and chronic inflammation, organ
 CC transplant rejection, Graft-versus-host disease (GVHD), lymphoid cell
 CC malignancies, septic and other forms of shock, loss of immune

CC responsiveness (as seen in human immunodeficiency virus (HIV)
CC infections) and failure of the immune response to tumour growth.
XX
SQ Sequence 225 AA; 70.6%; Score 1020; DB 21; Length 225;
Query Match 88.8%; Pred. No. 2.9e-89;
Best Local Similarity 9; Mismatches 16; Indels 0; Gaps 0;
Matches 199; Conservative 9; Indels 0; Gaps 0;
OY 61 LIGLALACGLLAVVSGSRASIAOEPAOEELVAEEDODPSLEINPQTEESODAPPL 120
DB 2 LGLGLALACGLLAVVSGSRASIAOEPAOEELVAEEDODPSLEINPQTEESODAPPL 61
OY 121 NBLVPRRSAPKGRKTRARRARAAIAHYEVHPRPGODAGVDTVSGWEARINSSSPPLR 180
DB 62 EQLVPRRSAPKGRKTRARRARAAIAHYEVHPRPGODAGVDTVSGWEARINSSSPPLR 121
OY 181 YNRQIGEFIVTRAGLYLYLCQVHFDEGKAVYKLDLLVGVLAALRCLEEFSAATASSLCP 240
DB 122 YDROIGEFIVIRAGLYLYLCQVHFDEGKAVYKLDLLVGVLAALRCLEEFSAATASSPGR 181
OY 241 QRLCOVSGLLALRPGSSLRIRTLRPAHLKAAFLTYFGLFOVH 284
DB 182 QRLCOVSGLLALRPGSSLRIRTLRPAHLKAAFLTYFGLFOVH 225
RESULT 12
AAW93591
ID AAW93591 standard; Protein: 211 AA.
XX
AC AAW93591;
XX
DT 18-JUN-1999 (first entry)
XX
DE Mouse TNRL3 protein.
XX
KM Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
KM developmental abnormality; gestational abnormality; prostate cancer;
KM APO4; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
KM cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
KM apoptosis; mouse.
XX
OS Mus sp.
XX
PM WO9911791-A2.
XX
PD 11-MAR-1999.
XX
PF 04-SEP-1998; 98WO-US18393.
XX
PR 05-SEP-1997; 97US-0924634.
XX
PA (UNIM) UNIV WASHINGTON.
XX
PI Chaudhary PM;
XX
DR WPI: 1999-205191/17.
XX
N-PSDB; AAX23425.
XX
PT New Tumour Necrosis Factor family receptor polypeptides and ligands -
PT useful for diagnosis and treatment of prostate cancer and
PT developmental or gestational abnormalities
XX
PS Claim 40; Fig 13b; 156pp; English.
XX
CC This invention describes isolated Tumour Necrosis factor (TNF) family
CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
CC their active fragments. APO4 is useful for diagnosing prostate cancer
CC by determining levels of APO4 in an individual. Prostate cancer can also
CC be treated using APO4 selective binding agents linked to a therapeutic
CC moiety. APO4 polypeptides are also useful for identifying selective
CC binding agents, useful in diagnosis/treatment of disease by binding of

CC agents to the polypeptide/active fragment which is extracellular, or
CC expressed on the cell surface. The binding is preferably performed in
CC vivo. APO4 polypeptides/active fragments are also useful for screening
CC for agonists and antagonists by binding and observing the change in APO4
CC activity. Effective pharmacological agents useful in diagnosis or
CC treatment of disease are also identified using APO4 polypeptides/active
CC fragments and APO4 signal transducer molecules that specifically interact
CC with a cytoplasmic domain of APO4 and detecting a change in level of APO4
CC activity. The method is performed in vivo or in vitro. APO4 polypeptides
CC are all useful as immunogens for preparing antibodies. APO4 is also
CC useful for diagnosis/treatment of developmental or gestational
CC abnormalities. APO8 was transfect to human breast carcinoma cell line
CC MCF-7, and induced apoptosis.
XX
SQ Sequence 211 AA; 67.0%; Score 968; DB 20; Length 211;
Query Match 89.1%; Pred. No. 2.5e-84;
Best Local Similarity 9; Mismatches 14; Indels 0; Gaps 0;
Matches 188; Conservative 9; Indels 0; Gaps 0;
OY 74 LNVVSGSRASIAOEPAOEELVAEEDODPSLEINPQTEESODAPPLNRLVPRRSAPKG 133
DB 1 LNVVSGSRASIAOEPAOEELVAEEDODPSLEINPQTEESODAPPLNRLVPRRSAPKG 60
OY 134 RKRARRAAIAHYEVHPRPGODAGVDTVSGWEARINSSSPPLRYNRQIGEFIVTRA 193
DB 61 RKRARRAAIAHYEVHPRPGODAGVDTVSGWEARINSSSPPLRYNRQIGEFIVTRA 120
OY 194 GLYLYLCQVHFDEGKAVYKLDLLVGVLAALRCLEEFSAATASSIPOLRCOVSGLLAL 253
DB 121 GLYLYLCQVHFDEGKAVYKLDLLVGVLAALRCLEEFSAATASSIPOLRCOVSGLLAL 180
OY 254 RPESSLRIRTLRPAHLKAAFLTYFGLFOVH 284
DB 181 RPESSLRIRTLRPAHLKAAFLTYFGLFOVH 211
RESULT 13
AAW29746
ID AAW29746 standard; Protein: 189 AA.
XX
AC AAW29746;
XX
DT 27-OCT-1998 (first entry)
XX
DE TNF related endothelium proliferative agent protein 2.
XX
KM TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
KM tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
XX
OS Homo sapiens.
XX
PM WO9835061-A2.
XX
PD 13-AUG-1998.
XX
PF 12-FEB-1998; 98WO-US02859.
XX
PR 10-FEB-1998; 98US-0021706.
XX
PR 12-FEB-1997; 97US-0798692.
XX
PA (ABBO) ABBOTT LAB.
XX
PI Wiley SR;
XX
DR WPI: 1998-447255/38.
XX
PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
PT treatment of autoimmune disease, tumours and inflammation
XX
PS Claim 16; Page 125-6, 142pp; English.
XX
CC The TNF-related endothelium proliferative agent (TREPA), or its

CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
CC promote wound healing or tissue grafting, by promoting vascularisation,
CC also to induce apoptosis for treating cancer and eliminating autoreactive
CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
CC TREPA peptides can also be used to target cytotoxic agents or for
CC affinity isolation of the corresponding receptor, the nucleic acid for
CC which can be used to transform tumour cells to render them more
CC responsive to TREPA and to screen for TREPA mimics.
CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat
CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
CC vascularisation), inflammation or a wide range of autoimmune conditions,
CC conditions involving abnormal stimulation of epithelial cells (e.g.
CC atherosclerosis), for birth control (inhibiting ovulation and placental
CC formation) or other angiogenic conditions (e.g. ulcers).
CC
XX
SO Sequence 189 AA;
Query Match 54.8%; Score 792; DB 19; Length 189;
Best Local Similarity 99.3%; Pred. No. 1.4e-67;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 132 KGRTRARRAIAHYEVHPRGODGAGVGTGSGWEARINSSPLRYNRQIGEFIVT 191
DB 37 KGRTRARRGIAAHYEVHPRGODGAGVGTGSGWEARINSSPLRYNRQIGEFIVT 96
OY 192 RAGLYLYICQVHFDEGKAVALKLDLVGVALALRCLEEFSAATASSLGPQLRCOVSGLL 251
DB 97 RAGLYLYICQVHFDEGKAVALKLDLVGVALALRCLEEFSAATASSLGPQLRCOVSGLL 156
OY 252 ALRPGSSLRIRTLPMALHKAAPFLTYFGLFOVH 284
DB 157 ALRPGSSLRIRTLPMALHKAAPFLTYFGLFOVH 189
RESULT 14
ID AAE00892 standard; Protein: 189 AA.
AC AAE00892;
DE 04-JUL-2001 (first entry)
DT 04-JUL-2001 (first entry)
XX
XX Human U14flag TREPA soluble construct.
DE
XX Human: tumour necrosis factor; TNF; angiogenesis; wound healing; tumour;
KM TREPA; TNF related endothelium proliferative agent; metastasis; grafting;
KW vulnery; HUVEC; human umbilical vein endothelial cell; U14flag.
OS
XX Homo sapiens.
OS
XX
PN US6207642-B1.
XX
XX 27-MAR-2001.
PD
XX
XX 26-JUN-1998; 98US-0105343.
PF
XX 12-FEB-1997; 97US-0798692.
PR 10-FEB-1998; 98US-0021706.
XX
XX (ABBO) ABBOTT LAB.
PA
XX
XX WILEY SR;
PI
XX
XX WPI: 2001-280760/29.
DR
XX
XX
XX Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
XX
XX Example 2; Column 75-78; 53pp; English.
XX
XX The present invention relates to extracellular signal molecules,

CC particularly members of tumour necrosis factor (TNF) family molecules
CC designated as TREPA (TNF related endothelium proliferative agent).
CC Soluble biologically active TREPA are used to treat TREPA-associated
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
CC in human for promoting wound healing and for vascularising grafted
CC tissue for successful grafting and to promote tissue grafts. The present
CC amino acid sequence is human U14flag TREPA soluble construct. This
CC sequence which is a biologically active molecule is capable of inducing
CC proliferation in HUVEC (human umbilical vein endothelial cells) cells.
XX
XX
SO Sequence 189 AA;
Query Match 54.8%; Score 792; DB 22; Length 189;
Best Local Similarity 99.3%; Pred. No. 1.4e-67;
Matches 152; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
OY 132 KGRTRARRAIAHYEVHPRGODGAGVGTGSGWEARINSSPLRYNRQIGEFIVT 191
DB 37 KGRTRARRGIAAHYEVHPRGODGAGVGTGSGWEARINSSPLRYNRQIGEFIVT 96
OY 192 RAGLYLYICQVHFDEGKAVALKLDLVGVALALRCLEEFSAATASSLGPQLRCOVSGLL 251
DB 97 RAGLYLYICQVHFDEGKAVALKLDLVGVALALRCLEEFSAATASSLGPQLRCOVSGLL 156
OY 252 ALRPGSSLRIRTLPMALHKAAPFLTYFGLFOVH 284
DB 157 ALRPGSSLRIRTLPMALHKAAPFLTYFGLFOVH 189
RESULT 15
ID AAE00895 standard; Protein: 146 AA.
AC AAE00895;
DE 04-JUL-2001 (first entry)
DT 04-JUL-2001 (first entry)
XX
XX Human TREPA (TNF related endothelium proliferative agent) fragment.
DE
XX Human: tumour necrosis factor; TNF; angiogenesis; wound healing;
KM TREPA; TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulnery.
OS
XX Homo sapiens.
OS
XX
PN US6207642-B1.
XX
XX 27-MAR-2001.
PD
XX
XX 26-JUN-1998; 98US-0105343.
PF
XX 12-FEB-1997; 97US-0798692.
PR 10-FEB-1998; 98US-0021706.
XX
XX (ABBO) ABBOTT LAB.
PA
XX
XX WILEY SR;
PI
XX
XX WPI: 2001-280760/29.
DR
XX
XX
XX Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
XX
XX
XX Example 14; Fig 1; 53pp; English.
XX
XX The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molecules
CC designated as TREPA (TNF related endothelium proliferative agent).
CC Soluble biologically active TREPA are used to treat TREPA-associated
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
CC in human for promoting wound healing and for vascularising grafted
CC tissue for successful grafting and to promote tissue grafts. The present
CC amino acid sequence is human U14flag TREPA soluble construct. This
CC sequence which is a biologically active molecule is capable of inducing
CC proliferation in HUVEC (human umbilical vein endothelial cells) cells.
XX
XX
XX
XX Inducing angiogenesis in mammal at desired sites for promoting wound
PT healing, by administering soluble fragment of extracellular domain of
PT tumor necrosis factor related endothelium proliferative agent protein
PT
XX
XX
XX Example 14; Fig 1; 53pp; English.
XX
XX The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molecules
CC designated as TREPA (TNF related endothelium proliferative agent).
CC Soluble biologically active TREPA are used to treat TREPA-associated
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
CC in human for promoting wound healing and for vascularising grafted

CC tissue for successful grafting and to promote tissue grafts.
 CC The present amino acid sequence is human TREPA fragment.

XX
 SQ Sequence 146 Aa;

Query Match 52.7%; Score 761; DB 22; Length 146;
 Best Local Similarity 100.0%; Pred. No. 9.3e-65;
 Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 139 RRAIAAHVEVHPPRGODGAQAGVDGTVSGWEARINSSPLRYNROIGETVTRAGLYYL 198
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 1 RRAIAAHVEVHPPRGODGAQAGVDGTVSGWEARINSSPLRYNROIGETVTRAGLYYL 60
 OY 199 YCQVHFDEGKAVYLLKLDLLVDGVLALRCLLEFSATAASSLGQRLCQVSGLLALRPGSS 258
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 61 YCQVHFDEGKAVYLLKLDLLVDGVLALRCLLEFSATAASSLGQRLCQVSGLLALRPGSS 120
 OY 259 LRIRTLPMWHLKAPFLTYFGLFOVH 284
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 121 LRIRTLPMWHLKAPFLTYFGLFOVH 146

Search completed: March 31, 2003, 07:44:12
 Job time : 55.5639 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 00:53:03 ; Search time 3581.52 Seconds
(without alignments)
2307.731 Million cell updates/sec

Title: US-09-245-198a-4
Perfect score: 1444
Sequence: 1 MSLLDFEISARRLPLPRSLG.....PMAHLKAAPLTYGFLQVH 284

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPRO/spool/US09245198/tunat_24032003.163553.25950/app.query.fasta.1.846
-DB=genbml -QMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCLALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09245198.acgn.1.1.3077.ernat.24032003.163553.25950 -NCPU=6 -ICPU=3
-NO.XLPXY -NO.MMAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Genbml:*
1: gb_ba:*
2: gb_hlg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*

29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pla:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1366	94.6	1353	6	AX201324
2	1366	94.6	1368	9	AF055872
3	1331.5	92.2	1651	9	BC019047
4	1298	89.9	1306	9	AF030099
5	1265	87.6	1236	6	ARI40407
6	1066	73.8	898	6	AX180714
7	1020	70.6	1168	10	AF030100
8	589.5	40.8	177703	2	AC016876
9	589.5	40.8	215795	2	AC127470
10	557	38.6	177555	2	AC130192
11	555.5	38.5	161428	2	AC126925
12	555.5	38.5	203083	2	AC069459
13	555.5	38.5	234182	10	AL603707
14	495	34.3	153553	2	AC126921
15	489.5	33.9	138792	2	AC119115
16	385.5	26.7	184026	2	AC098923
17	214	10.1	203281	2	AC126237
18	145.5	9.7	187835	8	AP002482
19	140	9.7	42210	1	SC1C
20	138.5	9.6	190050	1	AL646059
21	136.5	9.5	242930	2	AC105470
22	135	9.3	6397	1	AF002222
23	135	9.3	10728	1	AE004651
24	134	9.3	10965	1	AE004740
25	130	9.0	40352	1	SC61A
26	129	8.9	5751	1	SGDNASTRS
27	129	8.9	16185	1	SCC46
28	129	8.9	25459	1	SGAJ6985
29	129	8.9	32941	1	SCM10
30	129	8.9	43961	1	SCM10
31	128	8.9	137833	9	CNS01DV6
32	127	8.8	194304	9	AC016708
33	126.5	8.8	207050	9	AL646063
34	125.5	8.7	999	9	HSBARD11
35	125.5	8.7	28795	1	AVINIRC
36	125.5	8.7	87144	8	AP003822
37	125.5	8.7	153595	2	AP004272
38	125	8.7	30590	1	SC4G2
39	125	8.7	191221	2	AP005292
40	124.5	8.6	36734	1	SC6G10
41	124.5	8.6	98504	2	AP003876
42	124.5	8.6	107132	2	AP003911
43	124.5	8.6	154746	14	HSV2HG52
44	124.5	8.6	176580	2	AP003634
45	124	8.6	3957	6	A45258

RESULT 1

ALIGNMENTS

AX201324	LOCUS	AX201324	Sequence 3 from Patent WO0153486.	1353 bp	DNA	linear	PAT 30-AUG-2001
DEFINITION	AX201324	AX201324	Sequence 3 from Patent WO0153486.	1353 bp	DNA	linear	PAT 30-AUG-2001
ACCESSION	AX201324	AX201324	Sequence 3 from Patent WO0153486.	1353 bp	DNA	linear	PAT 30-AUG-2001
VERSION	AX201324.1	GI:15391154					
KEYWORDS							
SOURCE	human.						
ORGANISM	Homo sapiens						
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.						
AUTHORS	1 (bases 1 to 1353)						
	Askenazi,A.J., Goddard,A., Godowski,P.J., Gunney,A.L., Hillan,K.J., Masters,S.A., Pan,J., Pittl,R.M., Roy,M.A., Smith,V...						
TITLE	Stone,D.M., Watanabe,C.K. and Wood,W.I. Compositions and methods for the treatment of tumour						
JOURNAL	Patent: WO 0153486-A 3 26-JUL-2001; Genentech, Inc. (US)						
FEATURES	source	1..1353	Location/Qualifiers				
		/organism="Homo sapiens"					
		/db_xref="taxon:9606"					
BASE COUNT	257 a	443 c	389 g	264 t			
ORIGIN							
Alignment Scores:							
Pred. No.:	4,66e-63	Length:	1353				
Score:	1366.00	Matches:	268				
Percent Similarity:	100.00%	Conservative:	0				
Best Local Similarity:	100.00%	Mismatches:	0				
Query Match:	94.60%	Indels:	0				
DB:	6	Gaps:	0				
US-09-245-198A-4 (1-284) x AX201324 (1-1353)							
QY	17	ATGserLeuGlySerArgAspGlyAlaValArgGlnAlaGlnProProAlaProMet	36				
DB	1	CGATCCCTCGGGTCCCGGGATGGGGGGCGTGAAGCAGCACACCCCGCCCATG	60				
QY	37	AlaAlaArgArgSerGlnArgArgArgGlyArgGlyProGlyThrAlaLeuLeu	56				
DB	61	GCGGCGCGCGAGCAGAGCGGAGGGGGCGCCCGGGGGAGCCGCGCACCGCCCTGCTG	120				
QY	57	ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal	76				
DB	121	GTCGCCCTCGCCTGGGCTGGGCTGGGCTGGGCTGGGCTGGCTGGCTGGCTGGCTGG	180				
QY	77	ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGlnProAlaGlnGlnLeuVal	96				
DB	181	GTCAGTTGGGGAGCGCGGCGCATGCTGTCCGCCAGAGACCTGCCAGAGAGCTGGT	240				
QY	97	AlaGlnGluAspGlnAspProSerGluLeuLeuAsnProGlnThrGluGluSerGlnAspPro	116				
DB	241	CGAGAGGAGAGACCAAGACCCGTCGGAGACTGAATCCCGACAGAGAAAGCCAGATCCT	300				
QY	117	AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThr	136				
DB	301	GCGCCTTTCCTGAACCGCACTAATTGGGCTCGCAAGATGGCACTTAAAGCGCGGAAACA	360				
QY	137	ArgAlaArgArgAlaIleAlaAlaIleIstArgIuValHisProArgProGlyGlnAspGly	156				
DB	361	CGGGCTCGAAGACCATGCGACGCCCATTTATGAATTCATCCAGCATGGAGACAGCGGA	420				
QY	157	AlaGlnAlaGlyValaAspGlyThrValSerGlyTTPGlnGlnAlaArgIleAsnSerSer	176				
DB	421	GCGCAGGAGGCTGCGAGCGGACATGAGTGGCTGGAGAGAACCCAGATCAACAGCTCC	480				
QY	177	SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr	196				
DB	481	AGCCCTCTGCCTACAAACCGCCAGATCGGGAGTTATATGACACCCGGGCGGCTTAC	540				
QY	197	TyrLeuTyrCysGlnValHisPheAspGlnGlyGlyAlaValTyrLeuLysLeuAspLeu	216				
DB	541	TACCGTACGTGACAGTGCACCTTGTATAGAGGGAAGGCTGTCTACTCTTAACCTGACATTG	600				

QY	217	LeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSer	236
Db	601	CGGTGGATGTGTCTGCTGGCCCTGGCTGGCGGAATTCACGACACACGCGGAGT	660
QY	237	SerLeuGlyProGlnLeuArgLeuGysGlnValSerGlyLeuLeuAlaLeuArgProGly	256
Db	661	TCCCTCGGGCCCCACCTCCGCTTCGCGACGTGTCTGGCGCTTGGCCCCCTCGGCGAAGG	720
QY	257	SerSerLeuArgGlyLeuArgThrLeuProTrpPheAlaHisLeuLysAlaAlaProPheLeuThr	276
Db	721	TCCCTCCGTGGGATCCGACCCCTCCGCGGGCCATCTCAAGGGCGGCCCTTCCTGCAC	780
QY	277	TyrPheGlyLeuPheGlnValHis	284
Db	781	TACTTCGGACTCTTCACAGTTCAC	804
RESULT 2			
AF055872			
LOCUS	AF055872	1368 bp	mRNA
DEFINITION	Homo sapiens Ap03/DR3 ligand (Ap03L) mRNA, complete cds.	linear	PRI 04-MAY-1998
ACCESSION	AF055872		
VERSION	AF055872.1	GI:3108230	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens.		
	Homo sapiens.		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
REFERENCE	1 (bases 1 to 1368)		
AUTHORS	Marsters,S.A., Sheridan,J.P., Pittli,R.M., Brush,J., Goddard,A. and Ashkenazi,A.		
TITLE	Identification of a ligand for the death-domain-containing receptor Ap03		
JOURNAL	Curr. Biol. 8 (9), 525-528 (1998)		
MEDLINE	98322835		
PUBMED	9560343		
REFERENCE	2 (bases 1 to 1368)		
AUTHORS	Marsters,S.A., Sheridan,J.P., Pittli,R.M., Brush,J., Goddard,A. and Ashkenazi,A.		
TITLE	Direct Submission		
JOURNAL	Submitted (25-MAR-1998) Molecular Oncology, Genentech, 1 DNA Way, South San Francisco, CA 94080, USA		
FEATURES			
source	Location/Qualifiers		
	1..1368		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="17"		
	/map="17p13"		
	1..1368		
	/gene="Ap03L"		
	58..807		
	/gene="Ap03L"		
	/function="binds to the death-domain-containing receptor Ap03/DR3"		
	/note="TNF homolog"		
	/codon_start=1		
	/product="Ap03/DR3 ligand"		
	/protein_id="AAC39724.1"		
	/db_xref="GI:3108231"		
	/translation="MAARRSQRRRGREGPTALLPVALGDLALACGLLAVVSLGSRSLSAQEPADQELVAEDDQSEINPQTESQDPAPLNLVPRRSAPKGRKTRAKRAIAHVEHPRPGDGAQAGDGVTSQMEBARINSSPLRYNRQIGFIVTRAGLYRILCOVHPDEGRKAVYKLIDLIVGVALGCLFEFSATVAASSIGPOLRICOVSGLTALRPQSLRIRLRLPMNHAKAAPFLIVYGLFGVH"		
BASE COUNT	272 a	443 c	389 g
ORIGIN			264 t
Alignment Scores:			
Pred. No.:	4	71e-63	Length: 1368
Score:	1366.00	Matches: 268	
Percent Similarity:	100.00%	Conservative: 0	
Best Local Similarity:	100.00%	Mismatches: 0	
Query Match:	94.60%	Indels: 0	

DB:	9	Gaps:	0
US-09-245-198a-4 (1-284) x AF055872 (1-1368)			
Oy	17	ArgSerLeuGlySerArgAspGlyAlaValArgGlnAlaGlnProProAlaProMet	36
Db	1	CGATCCCTCGGGGTCGCGGATGGGGGGGGGTGAGGAGCAGACGCCGCCGCCCATG	60
Oy	37	AlaAlaArgArgSerGlnArgArgArgGlyArgArgGlyGluProGlyThAlaLeuLeu	56
Db	61	GCCGCCCGGTGGAGCCGAGAGCGAGGGGGCGCGGGGGGAGCCGCGCCCTGTG	120
Oy	57	ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal	76
Db	121	GTCCCGCTCGCGCTGGGCTTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG	180
Oy	77	ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuVal	96
Db	181	GTCAGTTGGGAGCCCGGAGCATGCTGTGCGCCAGAGAGCTGCCAGAGAGAGCTGTG	240
Oy	97	AlaGluGlnAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspPro	116
Db	241	GCAGAGAGAGCCAGGAGCCCTCGGAATCGAACCCAGACAGAAAGAGCAGAGATCT	300
Oy	117	AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLysThr	136
Db	301	GCAGCTTTCCGAAACCCGACTAGTTGGGCTCGCAGAGAGTCACTTAAGGCCGGAAACA	360
Oy	137	ArgAlaArgArgAlaAlaAlaAlaHisTyrGluValHisProArgProGlyGlnAspGly	156
Db	361	CGGGCTCGAAGAGGATCGCAGCCATTATGAGTTTCACAGACCTCGAGAGAGCGA	420
Oy	157	AlaGlnAlaGlyValaLeuAspGlyThrValSerGlyTrrpGluAlaArgLysSerSer	176
Db	421	GCGAGGAGGTGTGGAGCGGAGAGTGTGTGGGAGGAGGAGGAGATCAACAGCTCC	480
Oy	177	SerProLeuArgTyrAsnArgGlnIleGlyLysIleValThrArgAlaGlyLeuTyr	196
Db	481	AGCCCTCGCCCTACAAACCCGCAATCGGGAGTTTATAGTACCCGGGCTGGGCTCTAC	540
Oy	197	TyrLeuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeu	216
Db	541	TACCTGACTGTGAGTGACATTTGATGAGGAGGAGGAGGCTGTACTGAACTGAGACTTG	600
Oy	217	LeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSer	236
Db	601	CTGTGTGATGT	660
Oy	237	SerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgProGly	256
Db	661	TCCCTCGGCGCCGAGCTCCGCTGTCCAGAGTGTGTGGGCTGTGGCCCTCGGCGCAGG	720
Oy	257	SerSerLeuArgIleArgThrLeuProTrrpAlaHisLeuLysAlaAlaProPheLeuThr	276
Db	721	TCCCTCGGCGAGATCCGACCCCTCCCTGGGCGCATTCAGAGCTGCGCCCTCTCTCAC	780
Oy	277	TyrPheGlyLeuPheGlnValHis	284
Db	781	TACTTCGAGCTCTTCAGGTTTAC	804

RESULT 3
BC019047 1651 bp mRNA linear PRI 11-DEC-2001

LOCUS
DEFINITION Homo sapiens, similar to tumor necrosis factor (ligand)
superfamily, member 12, clone MGC:20669 IMAGE:4766071, mRNA,
complete cds.

ACCESSION
VERSION BC019047
KEYWORDS MGC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE	1 (bases 1 to 1651)		
AUTHORS	Strausberg, R.		
TITLE	Direct Submission		
JOURNAL	Submitted (07-DEC-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA		
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov		
COMMENT	Contact: MGC help desk Email: cgaps-remail.nih.gov Tissue Procurement: Louis Staudt CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisc.nih.gov/nisc/mgcengr1.nih.gov Contact: nisc.mgcengr1.nih.gov Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R., Lim, M., Maduro, Q.L., Mastello, C., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Snyder, B., Stantip, S., Thomas, P.J., Tiongson, E.E., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Zhang, L.-H. and Green, E.D.		
FEATURES	Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Series: IRAL Plate: 30 Row: P Column: 5 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA g1: 4507596. Location/Qualifiers 1..1651 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="MGC:20669 IMAGE:4766071" /tissue_type="Primary B-Cells from Tonsils" /clone_id="NIH-MGC_48" /lab_host="DH10B-R" /note="Vector: pOTB7" 106..510 /codon_start=1 /product="Similar to tumor necrosis factor (ligand) superfamily, member 12" /protein_id="AAH19047.1" /db_xref="GI:17512139" /translation="MAARRSQRRRGRRGERTALVPLALGLIALACGLLAVSLGSRASLSAQPADQELVAEDQDPSLEINPOTERSODPAPPLNRLVRRRSAPRGKTRRRARRAIAHYEVHRRPGDGAQADGGYTCRLP" CDS BASE COUNT 344 a 517 c 481 g 309 t ORIGIN		
Alignment Scores:	Pred. No.: 3.53e-61 Length: 1651 Score: 1331.50 Matches: 275 Percent Similarity: 77.25% Conservative: 0 Best Local Similarity: 77.25% Mismatches: 1 Query Match: 92.21% Indels: 81 DB: 9 Gaps: 1		
US-09-245-198A-4 (1-284) x BC019047 (1-1651)			
Oy	9	SerAlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAlaValArg	28
Db	25	TCCGCCCGCCGCTCCCTCCCGATCCCTCGGATGGGGGGGGCGGTAGG	84
Oy	29	GlnAlaGlnProProAlaProMetAlaAlaAlaArgArgSerGlnArgArgGlyArgArg	48
Db	85	CAGGCACAGCCCCCGCCCATGGCCCGCTGGAGACCAAGAGCGAGGGCGCGCG	144
Oy	49	GlyGluProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuAlaLeuAla	68
Db	145	GGGAGCGCGGCGCGCTGTGCTCGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG	204

[illegible]

AUTHORS	Chicheportiche Y., Boudon, P. R., Xu, H., Hsu, Y. M., Scott, H., Hession, C., Garcia, I. and Browning, J. L.							
TITLE	TWEAK, a new secreted ligand in the tumor necrosis factor family that weakly induces apoptosis							
JOURNAL	J. Biol. Chem. 272 (51), 32401-32410 (1997)							
MEDLINE	98070415							
PUBMED	9405449							
REFERENCE	2 (bases 1 to 1306)							
AUTHORS	Boudon, P., Hession, C., Tizard, R. and Browning, J.							
TITLE	Direct Submission							
JOURNAL	Submitted (14-OCT-1997) Cell Biology, Biogen, 12 Cambridge Center, Cambridge, MA 02142, USA							
FEATURES	location/Qualifiers							
SOURCE	1. 1306							
	/organism="Homo sapiens"							
	/db_xref="taxon:9606"							
	/chromosome="17"							
	/map="17p13"							
	/tissue_type="tonsil"							
	18. 767							
	/note="Ligand in the TNF family; secreted protein; start codon not verified experimentally"							
	/codon_start=1							
	/product="TWEAK"							
	/protein_id="AAC51923.1"							
	/db_xref="GI:2707219"							
	/translation="MAARRSQRRRGRRGEPTALLVPLALGLALACLLLAVALSLSRSIAOEPAOEELVAEDODPELSNPTEESODPAPLNLVRPBRAPGKRRARRALIAEYHVRPGDDGAGVAGVSGEERARSINSPLRYNRIGETVYRGLLYLYCOVHPDEGKAVYKLLDNGVVALGLCEFSATMASSLSGLPOLRQVSGLLALRPSSLRKRTLPAAHLKAAPPLYTFGLFOYH"							
BASE COUNT	247 a	434 c	368 g	257 t				
ORIGIN								
Alignment Scores:								
Pred. No.:	1.53e-59	Length:	1306					
Score:	1298.00	Matches:	254					
Percent Similarity:	100.00%	Conservative:	0					
Best Local Similarity:	100.00%	Mismatch:	0					
Query Match:	89.89%	Indels:	0					
DB:	9	Gaps:	0					
US-09-245-198A-4 (1-284) x AF030099 (1-1306)								
QY	31	GINP	PROALA	PRMEFALAALAAARGARSGERGINARGARGLYARGAGLYGLU 50				
DB	3	CAGCC	CCCCGGCCCCCATGGCGCCGCTGGAGACAGAGCGGAGGGCGCGGGGGAG 62					
QY	51	PROGLY	THRALALEU	VALPROLEUAALAEUGLYLEUGLYLEUAALAEUCYSLEU 70				
DB	63	CGGGC	CACCGCCCTCTGTGTCCGCTCCGCTGGGCGCTGGCGCTGGCTGCTTC 122					
QY	71	GLYLEU	LEU	ALAVAI	VALSERLEUGLYSERARGALASERLEUSERALAGLUGLUPRO 90			
DB	123	GCCCT	CTGGCTGGCGCTGTGTCAGTTGGGGAGCCGGGCGATGCTGCCGCCAGAGAGCT 182					
QY	91	ALAGIN	GLUGLU	LEU	VALAIALGLUGLUSPGLINASPPROSERGLULEUASNPROGLINTHR 110			
DB	183	GCCCG	AGGAGAGCTGTGTGGCAGAGAGACACAGACCCGCTGGAACTGAATCCCCAGACA 242					
QY	111	GLUGLU	SERGINAS	PPROALA	PROPHLEUASNA	TRGLEUVALATGPROATARGISERALA 130		
DB	243	GAAAG	AACCCAGAGATCCCTGGCGCTTTCCTGAACGACATGATTCGGCCCGCAGAAAGTGCA 302					
QY	131	PROLYSG	LYARGLVSTH	ARGALAA	ARGARALALAEALAEALAHISTYRGGLVALIHISPRO 150			
DB	303	CCTAA	AGCGCCGAAAMACACGGGCTCGAAGACCGATCCGACGCCATTATGAAGTTCAATCCA 362					
QY	151	ARGP	PROGLYGINAS	PGLYALAGLALAGLYALVALSPGLYTHPRVALSERGLYTRPGLUGLU 170				
DB	363	CGACCT	GGACACAGGACGGAGCCAGGCTGTGGACGGGACAGTGAAGTGAAGTGGAGGAA 422					
QY	171	ALATAC	GLI	LEASN	SERSETSERPROLEU	ARRGTYRISAN	ARGGINILEGLYGLU	PHILEVAL 190

```

Db 423 GCCAGAACACAGCTCCAGCCCTCTGGCTACACACCGCCGAGATCGGGGACTTATATC 482
OY 191 ThrArgAlaGlyLeuYrYrLeuYrCysGlnValHisPheAspGluGlyAlaVal 210
Db 483 ACCGCGCTGGGCTCTACTACTACTACTACTACTACTACTACTACTACTACTACTACTACT 542
OY 211 TyrLeuYsLeuAspLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPhe 230
Db 543 TACTGAAAGCTGAGCTGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 602
OY 231 SerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuGlyCysGlnValSerGlyLeu 250
Db 603 TCAGCCACTGGGGCCAGTTCCTCGGGCCCGAGCTCGCCCTCGAGTGTCTGGGCTG 662
OY 251 LeuAlaLeuArgProGlySerSerLeuArgGlnLeuArgThrLeuProThrAlaHisLeuLys 270
Db 663 TTGGCCCTGGCGCCAGGGTCTCCCTCGGAGTCCGACCCCTCGGGCCCATCTCAG 722
OY 271 AlaAlaProPheLeuThrYrPheGlyLeuPheGlnValHis 284
Db 723 GCTGCCCTTCTCTCCTACTCTGAGACTTTCAGAGTTTAC 764

RESULT 5
ARI40407 1236 bp DNA linear PAT 16-JUN-2001
LOCUS Sequence 1 from patent US 6207642.
DEFINITION ARI40407
ACCESSION ARI40407
VERSION ARI40407.1 GI:14482903
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1236)
AUTHORS Willey,S.R.
TITLE Member of the TNF family useful for treatment and diagnosis of
JOURNAL Patent: US 6207642-A 1 27-MAR-2001;
FEATURES
source 1..1236
location/Qualifiers
BASE COUNT 225 a 416 c 358 g 237 t
ORIGIN

Alignment Scores:
Pred. No.: 7 Size=58 Length: 1236
Score: 1265.00 Matches: 248
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.60% Mismatches: 0
Query Match: 87.60% Indels: 0
DB: Gaps: 0

US-09-245-198a-4 (1-284) x ARI40407 (1-1236)
OY 36 MetaAlaAlaArgSerGlnArgArgArgGlyArgGlyGluProGlyThrAlaLeu 55
Db 1 ATGGCCGGCCGCTGGAGCAGACAGGAGGGGGCGGGGGAGCCGGCAGCCGCTG 60
OY 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75
Db 61 CTGTCCCGCGCGCTGGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTG 120
OY 76 ValValSerLeuGlySerArgAlaSerLeuSerLeuGlnGluProAlaGlnGluGlu 95
Db 121 GTGGTACGTTTGGGAGCGGGGCACTGCTGTCCGCCAGGAGCTTGCCAGAGAGCTG 180
OY 96 ValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAsp 115
Db 181 GTGGCAGAGGAGGAGCAGGAGCCGTCGGAATGATCCCGCAGACAGAAAGCAGAGAT 240
OY 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLys 135
Db 241 CTTGCGCTTCTCTGAAACGCACTAGTTCGGCTCGAAGAAAGTGCACCTAAAGCCGGA 300

```

```

OY 136 ThrArgAlaArgArgAlaAlaAlaHisGlyGluValHisProArgProGlyGlnAsp 155
Db 301 ACAGGCTCGAAGAGGATCGCAGCCATTATAGTTATCCAGCAGCAGCAGCAGCAGCAG 360
OY 156 GlyAlaGlnAlaGlyValAspGlyThrValSerGlyTrpGluGlnAlaArgLeuSer 175
Db 361 GGAGCGCAGGAGGTGTGTGAGCGGAGCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 420
OY 176 SerSerProLeuArgYrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195
Db 421 TCCAGCCCTGCGCCTCAACCCGAGATCGGAGATTATATAGCACCCGGCGGCTC 480
OY 196 TyrTyrrLeuYrCysGlnValHisPheAspGluGlyValAlaValTyrrLeuYsLeuAsp 215
Db 481 TACTACCTGTACTCTCAGGCTCCTTGTATAGGAGGAGGCTGTCTACTTGAACCTTGAC 540
OY 216 LeuLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAla 235
Db 541 TTGCTGGTGGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
OY 236 SerSerLeuGlyProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuAlaLeuArgPro 255
Db 601 AGTTCCTCGGGCCCGAGCTCGCCCTCGCAGGCTGTGTGGCTGTGTGGCTCGGCGCA 660
OY 256 GlySerSerLeuArgGlyLeuArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeu 275
Db 661 GGTCTCTCCCTCGGAGATCCGACCCCTCCCTCGGGCCCATCTCAAGGCTGCCCTTCTCTC 720
OY 276 ThrTyrrPheGlyLeuPheGlnValHis 284
Db 721 ACTTACTTCGAGCTCTCCAGGTTTAC 747

RESULT 6
AXI80714 898 bp DNA linear PAT 06-AUG-2001
LOCUS Sequence 1 from Patent WO0145730.
DEFINITION AXI80714
ACCESSION AXI80714
VERSION AXI80714.1 GI:15132570
KEYWORDS
SOURCE synthetic construct.
ORGANISM synthetic construct
REFERENCE 1 (bases 1 to 898)
AUTHORS Willey,S.R.
TITLE Tweak receptor
JOURNAL Patent: WO 0145730-A 1 28-JUN-2001;
FEATURES
source 1..898
location/Qualifiers
BASE COUNT 187 a 266 c 267 g 178 t
ORIGIN

Alignment Scores:
Pred. No.: 1 Size=47 Length: 898
Score: 1066.00 Matches: 207
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 73.82% Indels: 0
DB: Gaps: 0

CDS
52..873
/note="human TWEAK fusion protein construct"
/codon_start=1
/transl_table=1
/protein_id="CAC50004.1"
/db_xref="GI:15132571"
/translation="MATGRTSLLAFGLICLPWLOGSATSSDRMKQIEDKTEELLS
KIYHIENETARIKKLIGERTSSLSASASADPEAPQEEIVAEEDDPSELNPTQTES
QDAPPELNLVPRRSAPKGRKTRARAIAAHVEVPRPDQAGVDTVSGWBEA
RINSSPLRNQIGERIYTRAGLYYCOVHDEGKAVYLIKDLVDGVIALRLEE
FSATTAASSIGPOLRICOVSGLLALRPSSLRIRTLPMALHKAAPFLITGLPQVH"

```


Db 665 TTTCAGTTCCAC 676

RESULT 8
AC016876/c 177703 bp DNA linear HTG 06-AUG-2002

LOCUS AC016876.5 GI:21313830
DEFINITION Homo sapiens clone RP11-186B7, *** SEQUENCING IN PROGRESS ***, 10
unordered pieces.

AC016876
VERSION AC016876.5
KEYWORDS HTG; HTGS_PHASEL; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE human.

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, M.,
Baldwin, U., Barna, N., Beckery, R., Boguslavsky, L., Bouckgalter, B.,
Brown, A., Castle, A., Colangelo, M., Collins, S., Collumore, A.,
Cooke, P., DeRellano, K., Dewar, K., Domino, M., Donelan, L., Doyle, M.,
Ferreira, P., FitzHugh, W., Forrest, C., Funke, R., Gage, D.,
Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J., C. Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J.,
Lenczky, J., Liu, C., Locke, K., Macdonald, P., Marquis, N.,
Mcwan, P., McGurk, A., McKernan, K., McLaughlin, J., Meldrum, J.,
Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Tirrell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,
Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (08-DEC-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE
AUTHORS Birren, B., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouckgalter, B.,
Camarita, J., Chang, J., Chazaro, B., Choepel, Y., Collumore, A.,
Cook, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J., Dodge, S.,
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,
Gardyna, S., Gord, S., Graham, L., Grand-pierre, N., Hagos, B.,
Horton, L., Hulme, M., Iliev, I., Johnson, R., Jones, C., Kamat, A.,
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mienga, V.,
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,
O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,
Rohan, J., Roy, A., Schauer, S., Schnupack, R., Seaman, S., Severy, P.,
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,
Tesfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,
Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission

JOURNAL Submitted (06-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT On Jun 2, 2002 this sequence version replaced gi:15421898.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIDR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L3849
Center clone name: 186_B_7

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces

```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
*       1       48645: contig of 48645 bp in length
* 48646 48745: gap of 100 bp
* 48746 58277: contig of 9532 bp in length
* 58278 58377: gap of 100 bp
* 58378 65804: contig of 7427 bp in length
* 65805 65904: gap of 100 bp
* 65905 79793: contig of 13889 bp in length
* 79794 79893: gap of 100 bp
* 79894 99493: contig of 19600 bp in length
* 99494 99593: gap of 100 bp
* 99594 111049: contig of 11456 bp in length
* 111050 111149: gap of 100 bp
* 111150 125020: contig of 13871 bp in length
* 125021 125120: gap of 100 bp
* 125121 145109: contig of 19989 bp in length
* 145110 145209: gap of 100 bp
* 145210 169458: contig of 24245 bp in length
* 169459 169558: gap of 100 bp
* 169559 177703: contig of 8145 bp in length.
*
* Location/Qualifiers
*       1..177703
*   /organism="Homo sapiens"
*   /db_xref="taxon:9606"
*   /clone_11b="RP11-186B7"
*   /clone_11b="RP11-11 Human Male BAC"
*
BASE COUNT      42862 a 45967 c 46588 g 41254 t 1052 others
ORIGIN

```

Alignment Scores:			
Pred. No.:	1,396-20	Length:	177703
Score:	589.50	Matches:	125
Percent Similarity:	65.62%	Conservative:	1
Best Local Similarity:	65.10%	Mismatches:	0
Query Match:	40.82%	Indels:	66
DB:	2	Gaps:	1
US-09-245-198A-4 (1-284) x ACO16676 (1-177703)			
QY 159	AlaAlValaspolyThrValSerGlyTPrgluAlaArgIleasnSerSerPro	178	
Db 37636	TCAGGTGGACGGGACAGTGAAGTGGCTGGAGGAGAACCGATACAGCTCCAGCCT		
QY 179	LeuArqTYrAsnArqGlnIleGlyGluPheIleValThrArAlaGlyLeuTYrTYrLeu	198	
Db 37576	CTGGGCTACACACCGCCAGATCGGGGAGTTATATGTCACCCGGGCTGGGCTACTACTG	3751	
QY 199	TYrCYs-----	200	
Db 37516	TACGTGAGGTAAAGCCCATCTGGTCATGGGTAAAGCAGTAAGAGAGTGGCGAAGGT	3745	
QY 200	-----	200	
Db 37456	TTGCCAGAGAGTGGGGACAAAGCTACAGGCTGGAGAGGTGAGTTGGGTTGGGTGG	3739	
QY 200	-----	200	
Db 37396	ATGGGATGCCTCGTCGTGAGGAAATTGGAAATTGAGCGAGGCGAGGAGGCTCG	3733	
QY 201	-----GlnValHisPheAspGluGlyLysAlaValTYrLeuLY	213	
Db 37336	ACTCGGCTGTGTCCCAACCCAGGTGACCTTGTATGAGGGGAAGGCTGTACTGAA	3727	
QY 213	SleuAspLeuValAspGlyValLeuAlaLeuArqCysLeuGluGluPheSerAlaTh	233	
Db 37276	GCTGACCTTCTGT	3721	
QY 233	rAlaAlaSerSerLeuGlyProGlnLeuArqLeuGlyCysGlnValSerGlyLeuAlaLe	253	

Db 37126 TGGGGAGATTCCCTGGGGCCAGCTCCGCTCTGCGCAGTGTGCTGTGGCCCT 37157
Qy 253 uatprrgolyserSerleuargilearqthrlleuprotfralahlisleulysalaalapr 273
Db 37156 GCGGCAAGGCTCTCCCTGGGATCGCACCCCTCCCTGGGCCATCTCAAGCTGCCCC 37097
Qy 273 opheleuthrtyrphedyleuphegnvalhlis 284
Db 37096 CTTCCTCACTACTTCGACTCTTCAGGTTTAC 37063
RESULT 9
AC127470 215795 bp DNA linear HTG 17-JUL-2002
AC127470
LOCUS
DEFINITION
AC127470 215795 bp DNA linear HTG 17-JUL-2002
Pan troglodytes clone RP43-145D13, WORKING DRAFT SEQUENCE, 15
unordered pieces.
AC127470
AC127470.1 GI:21886866
VERSION
HTG: HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Chimpanzee.
ORGANISM
Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
1 (bases 1 to 215795)
AUTHORS
Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Sternberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C.,
Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J.,
Haghighi,P., Hansen,N., Ho,S.-L., Idol,J.R., Karlins,E., Lalic,P.,
Lee-Jin,S.-O., Legaspi,R., Maduro,O.L., Maduro,Y.B.,
Marquillas,E.H., Mastello,C., Maskeri,B., Mastrisan,S.D.,
McCloskey,J.C., McDowell,J., Pegurigan,C., Pearson,R.,
Portnoy,M.E., Prasad,A., Schueler,M.G., Stantrilop,S., Thomas,J.W.,
Thomas,P.J., Touchman,J.W., Tsurgoun,C., Vogt,J.L., Walker,M.A.,
Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.
NISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 215795)
TITLE
JOURNAL
REFERENCE
AUTHORS
JOURNAL
COMMENT
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@nih.gov
Project Information
Center project name: cms
Center clone name: 145D13
Summary Statistics
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 203929 bases at least Q40
Consensus quality: 206865 bases at least Q40
Consensus quality: 209181 bases at least Q20
Insert size: 190000; agarose-fp
Insert size: 214395; sum-of-contigs
Quality coverage: 7.30x in Q20 bases; agarose-fp
Quality coverage: 6.47x in Q20 bases; sum-of-contigs
NOTE: This is a 'working draft' sequence. It currently
consists of 15 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as
runs of N, but the exact sizes of the gaps are unknown.
This record will be updated with the finished sequence
as soon as it is available and the accession number will
be preserved.
1 3593: contig of 3593 bp in length
* 3594 3693: gap of unknown length
* 3694 8001: contig of 4308 bp in length
* 8002 8101: gap of unknown length

8102 15639: contig of 7538 bp in length
* 15640 15739: gap of unknown length
* 15740 24775: contig of 9036 bp in length
* 24776 24875: gap of unknown length
* 24876 30150: contig of 5275 bp in length
* 30151 30250: gap of unknown length
* 30251 39212: contig of 8962 bp in length
* 39213 39312: gap of unknown length
* 39313 49342: contig of 10030 bp in length
* 49343 49443: gap of unknown length
* 49443 60956: contig of 11514 bp in length
* 60957 61056: gap of unknown length
* 61057 74520: contig of 13464 bp in length
* 74521 74620: gap of unknown length
* 74621 89522: contig of 14902 bp in length
* 89523 89622: gap of unknown length
* 89623 105537: contig of 15915 bp in length
* 105338 105637: gap of unknown length
* 105638 121197: contig of 15560 bp in length
* 121198 121297: gap of unknown length
* 121298 138883: contig of 17586 bp in length
* 138884 138983: gap of unknown length
* 138984 166802: contig of 27819 bp in length
* 166803 166902: gap of unknown length
* 166903 215795: contig of 48893 bp in length.
Location/Qualifiers
1. 215795
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-145D13"
/clone_1fb="RP43"
1. 3593
/note="assembly-fragment"
misc_feature 3694 . 8001
/note="assembly-fragment"
misc_feature 8102 . 15639
/note="assembly-fragment"
misc_feature 15740 . 24775
/note="assembly-fragment"
misc_feature 24876 . 30150
/note="assembly-fragment"
misc_feature 30251 . 39212
/note="assembly-fragment"
misc_feature 39313 . 49342
/note="assembly-fragment"
misc_feature 49443 . 60956
/note="assembly-fragment"
misc_feature 61057 . 74520
/note="assembly-fragment"
misc_feature 74621 . 89522
/note="assembly-fragment"
misc_feature 89623 . 105537
/note="assembly-fragment"
misc_feature 105638 . 121197
/note="assembly-fragment"
misc_feature 121298 . 138883
/note="assembly-fragment"
misc_feature 138984 . 166802
/note="assembly-fragment"
misc_feature 166903 . 215795
/note="assembly-fragment"
BASE COUNT 52160 a 54859 c 53752 g 53513 t 1511 others
ORIGIN
Alignment Scores:
Pred. No.: 1 7e-20 Length: 215795
Score: 589.50 Matches: 125
Percent Similarity: 65.62% Conservative: 1
Best Local Similarity: 65.10% Mismatches: 0
Query Match: 40.82% Indels: 66
DB: 2 Gaps: 1
US-09-245-198a-4 (1-284) x AC127470 (1-215795)

[illegible]

```

Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc_zoehngri.nih.gov
----- Project Information -----
Center project name: dbj
Center clone name: 436K21
----- Summary Statistics -----
Sequencing vector: plasmid; n/a; 100% of reads
Chemistry: dye-terminator Big Dye, 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 164120 bases at least Q40
Consensus quality: 168085 bases at least Q30
Consensus quality: 170767 bases at least Q20
Insert size: 169000; agarose-fp
Insert size: 176655; sum-of-contigs
Quality coverage: 4.78x in Q20 bases; agarose-fp
Quality coverage: 4.57x in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 3460: contig of 3460 bp in length
* 3461 3560: gap of unknown length
* 3561 6869: contig of 3309 bp in length
* 6870 6969: gap of unknown length
* 6970 13465: contig of 6496 bp in length
* 13466 13565: gap of unknown length
* 13566 21301: contig of 7736 bp in length
* 21302 21401: gap of unknown length
* 21402 33719: contig of 12318 bp in length
* 33720 33819: gap of unknown length
* 33820 45054: contig of 11235 bp in length
* 45055 45154: gap of unknown length
* 45155 58724: contig of 13570 bp in length
* 58725 58825: gap of unknown length
* 58826 76670: contig of 17846 bp in length
* 76671 76770: gap of unknown length
* 76771 115165: contig of 38395 bp in length
* 115166 115265: gap of unknown length
* 115266 117555: contig of 62290 bp in length.
*
Location/Qualifiers
1..117555
/organism="Sus scrofa"
/db_xref="taxon:9623"
/clone="RP44-436K21"
/clone_1lb="RP44"
1..3460
/note="assembly-fragment"
misc_feature 3561..6869
/note="assembly-fragment"
misc_feature 6970..13465
/note="assembly-fragment"
misc_feature 13566..21301
/note="assembly-fragment"
misc_feature 21402..33719
/note="assembly-fragment"
misc_feature 33820..45054
/note="assembly-fragment"
misc_feature 45155..58724
/note="assembly-fragment"
clone_end:17
vector_side:left"
58825..76670
/note="assembly-fragment"
76771..115165
/note="assembly-fragment"
115266..117555
/note="assembly-fragment"

```



```

Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
-----
Project Information
Center project name: MAFO
Center clone name: RP23-168P5
-----
Summary Statistics
Sequencing vector: ML3; L08821
Chemistry: Dye-Primer Bodipy: 48% of reads
Chemistry: Dye-Terminator Big Dye: 52% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 212648 bases at least Q40
Consensus quality: 218902 bases at least Q30
Consensus quality: 222384 bases at least Q20
Estimated insert size: 210656; sum-of-ctrls estimation
Quality coverage: 7.2x in Q20 bases; agarose-1p estimation
Quality coverage: 7.2x in Q20 bases; sum-of-ctrls estimation
-----
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/centbank\_drift\_data.html)
* NOTE: This is a "working draft" sequence. It currently
* consists of 7 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 62152: contig of 62152 bp in length
* 62153
* 62252: gap of unknown length
* 62253
* 118772: contig of 55520 bp in length
* 118773
* 118872: gap of unknown length
* 118873
* 148824: contig of 30052 bp in length
* 148925
* 149024: gap of unknown length
* 149025
* 167231: contig of 18207 bp in length
* 167232
* 167331: gap of unknown length
* 167332
* 169807: gap of 22576 bp in length
* 169808
* 190007: gap of unknown length
* 190008
* 196537: contig of 6530 bp in length
* 196538
* 196637: gap of unknown length
* 196638
* 203083: contig of 6446 bp in length.
*
* Location/Qualifiers
*
* 1..203083
*     /organism="Mus musculus"
*     /db_xref="taxon:10090"
*     /chromosome="11"
*     /clone="RP23-168P5"
*
BASE COUNT      52662 a  49293 c  47892 g  52633 t      603 others
ORIGIN
-----
Alignment Scores:
Pred. No.:      9,31e-19      Length:      203083
Score:          555.50      Matches:      117
Percent Similarity: 67.22%      Conservative:  4
Best Local Similarity: 65.00%      Mismatches:   5
Query Match:     38.47%      Indels:      54
                        Gaps:      1
US-09-245-198A-4 (1-284) x AC069459 (1-203083)
Oy      159      AAlAGlyVAlAsPcGlyThYrVAlSeRgLyTrPgLuGluAlAryIlleAsnSeRSeRPro 178
Db      43305      TCAGCTGGGAGGAGGACAGTGTGGCTGGGGAAGAGACCAATCAACAGCTCCAGCCCT 43246
Oy      179      LeuAGTyTAAsARyGlnlIeGlyGluPheIlleValTThARAlAGlyLeuTyTyTleu 198
Db      43245      CTGGCCCTACGACGCCAGATTGGGGAAATTACAGTCATCAGGGGCTGGGCTCTACTACTCG 43186
Oy      199      TyTCys----- 200
Db      43185      TACTCTCAGGTAAAGCCCTGCGCTTCAGGGGACAGCAAGGCTAAAGGAGGAGAGGA 43126
Oy      200      ----- 200

```

```

43125 CTGGCAGCAGGAMTCGGTGGGAGGAGACCTCGGACATCAGGAAGAGTCCTGGCTT 43066
OY 201 -----GlnValHisPheAs 205
Db 43065 TCGATGAGGAGGAGCGAGAGGTGTGATTTGCTTGGCTCTCTCTGTTCAGGTCACCTTTTGA 43006
OY 205 pGluValysAlaVal1YrLeuLysLeuAspLeuLeuValaspGlyVal1LeuAlaLeuAr 225
Db 43005 TGAGGGAAAGGCTGTCTTACCTGAACTTGACTTGTGTGAACGGTGTGCTGGCCCTGCG 42946
OY 225 gCysLeuLugLupPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuCy 245
Db 42945 CTGGCTGGAAGATTCCTCAGCGACAGCAGCAAGCTCTCTCTGGGCCGCCAGCTCCGTTTGTG 42886
OY 245 sGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArglyLeaAgThrLeuPr 265
Db 42885 CCAGGTGTCTGGGCTGTGGTCCGCTGCGGCCAGAGGCTTCCCTTGGATCCGACCTCC 42826
OY 265 oTPrpAlaHisLeuLysAlaAlaProPheLeuThyrPheGlyLeuPheGlnValHis 284
Db 42825 CTGGGCTCATCTTAAAGCTGTGCCCTTCTCTAACCTGTTGGACTTTCAAAGTTTCAAC 42768
RESULT 13
AL603707/c AL603707 234182 bp DNA linear ROD 17-NOV-2001
LOCUS Mouse DNA sequence from clone RP23-422L16 on chromosome 11,
DEFINITION complete sequence.
ACCESSION AL603707
VERSION AL603707.5 GI:17017790
KEYWORDS HTG.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Eureleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 234182)
Pearce/A.
Direct Submission
Submitted (17-NOV-2001) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
On Nov 20, 2001 this sequence version replaced gi:16605765.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em: EMBL, SW:
SWISSPROT, Tr: TREMBL, Wp: WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep RP23-422L16 is
from the RPc1-23 Mouse PAC Library
constructed by the group of Pieter de Jong.
For further details see http://www.chori.org/bacpac/home.htm
VECTORS: pBAC3.6
This sequence is the entire insert of clone RP23-422L16.
FEATURES
Source
1..234182
/organism="Mus musculus"
/db_xref="taxon:10090"
/chromosome="11"
/clone="RP23-422L16"
/clone_11p="RPc1-23"
/complement(84050..84131)
misc.feature

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 00:47:43 ; Search time 294.043 Seconds
(without alignments)
2175.080 Million cell updates/sec

Title: US-09-245-198a-4
Perfect score: 1444
Sequence: 1 MSLLDFEISARRLPLPSRLG.....PMAHLKAPFLTYGFLQVH 284

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2185239 segs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODL=frame+ p2n.model -DEV=x1h
-DB=cgn2_1/USPTO.spool/US09245198/runat_24032003_163552_25944/app_query.fasta.1.846
-LOOEXT=0 -UNITs-bits -START=1 -END=-1 -MATRIX=blomsum2 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MITLEN=0 -MAXLEN=200000000
-USER=US09245198.ecgn.1.1-338_etunal_24032003_163552_25944 -NCPU=6 -ICPU=3
-NO_XLPPY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMEOUT=120
-WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FPGAPOP=6 -FPGAEXT=7
-YGAPOP=10 -YGAEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the distribution.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1444	100.0	1373	19 AAV18600	Homo sapiens tumou
2	1406	97.4	1364	24 ABK34881	Human cDNA encodin
3	1377	95.4	1421	20 AAX56000	Human tumour necro
4	1366	94.6	1353	21 AAA49717	Human PRO207 cDNA
5	1366	94.6	1353	24 ABK40255	cDNA encoding huma
6	1265	87.6	1236	19 AAV47613	TNF related endothe
7	1265	87.6	1236	22 AAD04350	Human TREPA (TNF r
8	1066	73.8	898	22 AAS03964	Expression vector
9	1062	73.5	1030	22 AAX33424	Human TNRL3 DNA.
10	1020	70.6	1168	19 AAV18599	Mus musculus tumou
11	968	67.0	701	20 AAX23425	Mouse TNRL3 DNA.
12	127	8.8	2856	19 AAV24140	Homo sapiens BARD1
13	126.5	8.8	1660	21 AAA49205	Corn putative lecl
14	124.5	8.6	2832	20 AAV69289	Marine JNK-Interac
15	124.5	8.6	154746	24 AAD25519	Human herpesvirus
16	124	8.6	3957	22 AAA09686	HSV-2 immediate ea
17	124	8.6	154746	24 AAD25519	Human herpesvirus
18	123.5	8.6	11705	19 AAV62160	HSV-2 strain SB5 C
19	123.5	8.3	117213	19 AAV62176	CDNA of a human di
20	120.5	8.3	2828	24 AAA46674	Human benign prost
21	120.5	8.3	2828	24 ABK64503	Colon adenocarcino
22	120.5	8.3	2828	24 ABL61804	Thyroid cancer rel
23	117.5	8.1	3880	24 ABL67211	Human TATA-binding
24	117.5	8.1	3901	17 AAT42224	TATA-binding prote
25	117.5	8.1	3901	18 AAT79598	TATA-binding prote
26	117.5	8.1	3902	15 AAO70738	DNA encoding an ac
27	117	8.1	1030	20 AAV99230	Nucleotide sequenc
28	117	8.1	1030	21 AAT72514	Drosophila melanog
29	116	8.0	978	23 ABL21473	DNA encoding tumou
30	116	8.0	2148	24 ABK11680	Chicken beta-actin
31	115	8.0	1275	11 AAO01620	Expression vector
32	115	8.0	1344	12 AAO13318	Hybrid promoter of
33	115	8.0	1495	22 AAF84455	Chicken beta-actin
34	115	8.0	1495	22 AAF84455	Mouse Meg1/Grb10 n
35	115	8.0	4211	22 AAH21793	Rat expression pro
36	115	8.0	5504	24 AAL43172	Human expression p
37	115	8.0	5556	22 AAL43171	Human gene containi
38	115	8.0	5581	22 AAL41035	Positive selection
39	115	8.0	5759	24 ABK49521	PCAGGS-HSEPOR2 rat
40	115	8.0	6148	24 ABK49520	Human p531-alpha-
41	115	8.0	6256	24 AAL43173	HSV-2 strain SB5 C
42	115	7.9	8033	22 AAD12739	vitamin D hydroxyl
43	114.5	7.9	12700	19 AAV62133	Platenolide syntha
44	113.5	7.9	3836	15 AAO53997	
45	113.5	7.9	44377	18 AAT78508	

ALIGNMENTS

RESULT 1
AAV18600
AAV18600 standard; cDNA: 1373 BP.

AAV18600;
21-JUL-1998 (first entry)

DE Homo sapiens tumour necrosis factor related ligand (TRELL) gene.

XX TRELL: tumour necrosis factor related ligand; tnfr; treatment;

KW cancer; autoimmune disease; immune system; stimulation; suppression;

XX graft rejection; ds.

OS Homo sapiens.

XX Key Location/Qualifiers
FH CDS 1..852

```
FT      /*tag-a
XX      /note- "tumour necrosis factor related ligand"
PN      MO9805783-A1.
XX      12-FEB-1998.
PD      07-AUG-1997; 97WO-US13945.
XX      18-MAR-1997; 97US-0040820.
PR      07-AUG-1996; 96US-0023541.
PR      18-OCT-1996; 96US-0028515.
XX      (BIOJ ) BIOGEN INC.
PA      (UYGE-) UNIV GENEVA FACULTY MEDICINE.
XX      Browning JL, Chicheportiche Y;
PI      WPI; 1998-145619/13.
DR      P-PSDB; AAW47525.
XX      Tumour necrosis factor related ligand - useful for, e.g. treating
PT      cancer, auto-immune disease and immune responses to tissue grafts
XX      Claim 2; Pages 48-50; 69pp; English.
XX      The sequence is that encoding human tumour necrosis factor related
CC      ligand (TRELL). TRELL or active fragments can be included with a
CC      carrier in pharmaceutical compositions to treat cancer, autoimmune
CC      diseases or immune responses to tissue grafts, or to stimulate or
CC      suppress the immune system. It is useful to screen for TRELL
CC      receptors, by labelling with a detectable label and screening
CC      compositions for binding. Agents interfering with TRELL-receptor
CC      binding can also be screened for, can then be administered,
CC      optionally with interferon- gamma, to induce cell death or
CC      treat, suppress or alter immune responses (especially involving human
CC      adenocarcinoma cells) involving a signal pathway between TRELL and its
CC      receptor. The DNA sequence can be used in gene therapy for
CC      TRELL-related disorders in mammals (especially humans), e.g. tumours,
CC      autoimmune and inflammatory diseases or inherited genetic disorders,
CC      by introducing into cells, and expressing, therapeutically effective
CC      amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
CC      It may also be of use in the preparation of prepare probes for
CC      screening natural/synthetic DNAs for TRELL-encoding sequences
CC      and for antisense therapy.
XX      SQ      Sequence 1373 BP; 247 A; 462 C; 394 G; 270 T; 0 other;
Alignment Scores:
Pred. No.:      9,06e-96      Length:      1373
Score:      1444.00      Matches:      284
Percent Similarity:      100.00%      Conservative:      0
Best Local Similarity:      100.00%      Mismatches:      0
Query Match:      100.00%      Indels:      0
DB:      Gaps:      0
US-09-245-198a-4 (1-284) x AAV18600 (1-1373)
QY      1 MetSerLeuLeuAspPheGluIleSerAlaArgArgLeuProLeuProArgSerLeuGly 20
DB      1 ATGTCAATTCTTAACTTTGCAAAATTTCCGCCCGGCTCCCTCCCGCATTCCTCGG 60
QY      21 SerArgAspGlyGlyAlaValArgGlnAlaGlnProProAlaProMetLeuAlaAlaArg 40
DB      61 TCCCGGGATGGGGGGCGGTGAAGGACAGCCACCCCGCCCAATGGCGCCCTCGG 120
QY      41 SerGlnArgArgArgGlyArgArgGlyGluProGlyThrAlaLeuLeuValProLeuAla 60
DB      121 ACCCAAGAGCGGAGGGGGCGCGGGGAGCCGGGACCGCCTGCTGCTCCCGCTCGCG 180
QY      61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValAlaSerLeuGly 80
DB      181 CTGGGCTGGGCTGGGCTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGG 240
```

```
QY      81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluLeuValAlaGluGluAsp 100
DB      241 AGCCGGGCAATCGCTGTCCGCCAGAGAGCTGCCAGAGAGAGCTGTGGCAGAGAGAGAC 300
QY      101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeu 120
DB      301 CAGGACCCGCTCGGCACTGAATCCCGAGACAGAAAGGAGAGATCTCGCTTTCTCG 360
QY      121 AsnArgLeuValArgProArgSerAlaProLysGlyArgLysThrArgAlaArg 140
DB      361 AACCGACTAGTTGGCTTGCAGAACTGCACCTAAAGGCCGGAACACAGGCTCCAA 420
QY      141 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160
DB      421 GCGATCGCAGACCATTTATGAAGTTTCATCCAGACACCTGAGACAGCAGCAGCAGAG 480
QY      161 ValAspGlyThrValSerGlyTrpGluGluAlaArgIleAsnSerSerProLeuArg 180
DB      481 GTGGAGCGGACACTGAGTGGCTGGGAGAGAACCCAGATCAACAGCTCCAGCCCTCGCC 540
QY      181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200
DB      541 TACACCCGCGAGTCGGGAGTTTATAGTCACCCGGGCTGGCTTACTTACTTCTGT 600
QY      201 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220
DB      601 CAGGTCACTTGTATGAGGAGGAGGAGGCTGTCTACTGACCTGAGCTGTGCTGTG 660
QY      221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerLeuGlyPro 240
DB      661 GTGCTGGCCCTGGCGCTGGCGGAAATTCACGCCACCTGGGCCGAGTCCCTCGGGCCC 720
QY      241 GlnLeuArgLeuCysGlnValSerGlyLeuAlaLeuAlaLeuArgProGlySerSerLeuArg 260
DB      721 CAGCTCGGCTGTGCGCAGAGTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG 780
QY      261 IleArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 280
DB      781 ATCCGACACCTCCCGCGGCGCATCTCAAGAGGCGCCCTTCTCTCACAATTCGAGATC 840
QY      281 PheGlnValHis 284
DB      841 TTCACAGGTTTAC 852
RESULT 2
ID      ABRK34881
XX      ABRK34881 standard; cDNA; 1364 BP.
XX      ABRK34881;
AC      08-MAY-2002 (first entry)
DT      XX
XX      Human cDNA encoding secreted protein #19.
DE      Human: secreted protein; gene; ss; nutritional supplement; haemophilia;
XX      viral infection; bacterial infection; fungal infection; diabetes; asthma;
XX      autoimmune disorder; rheumatoid arthritis; multiple sclerosis; tumor;
XX      autoimmune thyroiditis; allergic reaction; neurodegenerative disease;
XX      Alzheimer's disease; Parkinson's disease; liver fibrosis; cancer; ulcer;
XX      coagulation disorder; inflammatory disorder; Crohn's disease; incision;
XX      tissue regeneration; wound healing; burn; hematopoiesis;
XX      myeloid cell deficiency; lymphoid cell deficiency.
OS      Homo sapiens.
XX      WO200177288-A2.
XX      18-OCT-2001.
XX      29-MAR-2001; 2001WO-US10224.
PF      06-APR-2000; 2000US-195582P.
XX      PR
```


CC to induce JNK/SAPK-dependent responses in mammalian cells.

XX Sequence 1421 BP; 281 A; 464 C; 272 T; 0 other;

Alignment Scores:

Pred. No.:	6,69e-91	Length:	1421
Score:	1377.00	Matches:	272
Percent Similarity:	97.84%	Conservative:	0
Best Local Similarity:	97.84%	Mismatches:	6
Query Match:	95.36%	Indels:	0
DB:	20	Gaps:	0

US-09-245-198a-4 (1-284) x AAX56000 (1-1421)

```

OY 7 GLUILESerAlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyAla 26
    |||||
DB 5 GAGATCCCTCGACCTGACCCAGCGCGTCCGCGATCCCTCGGGTCCCGGATGGGGGGCG 64

OY 27 ValArgGlnAlaGlnProProAlaPrometAlaAlaArgArgSerGlnArgArgGly 46
    |||||
DB 65 GTGAGCGAGGACAGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCG 124

OY 47 ArgArgGlyGlnProGlyThrAlaLeuLeuValProLeuAlaLeuGlyLeuAla 66
    |||||
DB 125 CCGCGGGGGAGCGGGGACCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 184

OY 67 LeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSerLeuSer 86
    |||||
DB 185 CTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 244

OY 87 AlaGlnGlnProAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 106
    |||||
DB 245 GCCCAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 304

OY 107 AsnProGlnThrGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 126
    |||||
DB 305 AATCCCGACAGCAAGAACCCAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 364

OY 127 ArgArgSerAlaProGlySerGlyArgGlyArgGlyArgGlyArgGlyArgGlyArg 146
    |||||
DB 365 CGCAGAGAGTGCACCTTAAGGCGCGAAACCGGGCTCGAAGAGCGATCGAGCGCATTTAT 424

OY 147 GlnValHisProArgProGlyGlnAspGlyAlaGlnAlaGlnAlaGlnAlaGlnAla 166
    |||||
DB 425 GAAGTTCATCCAGACCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 484

OY 167 GATTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 186
    |||||
DB 485 GCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 544

OY 187 GluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGlu 206
    |||||
DB 545 GAGTTTATAGTACCCGGGGGTGGGCTCTACTACTGCTACTGCTACTGCTACTGCTACTG 604

OY 207 GlyAlaAlaValTyrLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 226
    |||||
DB 605 GGGAGAGCTGTCTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 664

OY 227 LeuGlnGlnPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuGlyGln 246
    |||||
DB 665 CTGGAGGAAATTCACACCTGCGGAGTCTCTCTGCGGCCCGCCAGCTCTCTCTCTCTCA 724

OY 247 ValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgGlnLeuArgThrLeuProTrp 266
    |||||
DB 725 GGTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGT 784

OY 267 AlaHisLeuLeuAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
    |||||
DB 785 GCCCATCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 838

RESULT 4
AAX49717
ID AAX49717 standard: cDNA: 1353 BP.
  
```

```

XX AC AAX49717;
XX XX 25-SEP-2000 (first entry)
XX XX
XX DE Human PRO207 cDNA clone DNA30879-1152.
XX XX
XX KW PRO207; human; antitumour; tumour; therapy; cytostatic;
XX KW breast cancer; ovarian cancer; renal cancer; colorectal cancer;
XX KW uterine cancer; prostate cancer; lung cancer; bladder cancer;
XX KW central nervous system cancer; melanoma; leukaemia; neoplasm; ss.
XX OS Homo sapiens.
XX FH
XX FH Key Location/Qualifiers
XX FT CDS 58..807
XX FT sig_peptide 58..177
XX FT FT /tag= a
XX FT FT /tag= b
XX FT FT mat_peptide 178..804
XX FT FT /tag= c
XX XX
XX PN MO200037638-A2.
XX PD 29-JUN-2000.
XX PF
XX PF 02-DEC-1999: 99WO-US28565.
XX PR
XX PR 22-DEC-1998: 98US-0113296.
XX PR 08-MAR-1999: 99WO-US05028.
XX PR 21-APR-1999: 99US-0130232.
XX PR 28-APR-1999: 99US-0131445.
XX PR 14-MAY-1999: 99US-0134287.
XX PR 20-JUL-1999: 99US-0144758.
XX PR 26-JUL-1999: 99US-0145698.
XX PR 15-SEP-1999: 99WO-US21090.
XX PR 15-SEP-1999: 99WO-US21547.
XX PA
XX PA (GENTH ) GENENTECH INC.
XX PI
XX PI Ashkenazi AJ, Goddard A, Godowski PJ, Gurney AL, Marsters SA:
XX PI Napier MA, Pitti RM, Wood WI:
XX DR
XX DR WPI: 2000-442668/38.
XX DR P-PSDB: AAY95338.
XX PT
XX PT Novel composition to inhibit neoplastic cell growth or for treating
XX PT tumor in mammal comprises polypeptides PRO179, PRO307, PRO320, PRO219,
XX PT PRO221, PRO224, PRO328, PRO301, PRO362, PRO356, PRO309 or
XX PT PRO866
XX PS
XX PS Claim 20; Fig 3; 172pp; English.
XX XX
XX CC The present sequence is that of cDNA clone DNA30879-1152
XX CC (ATCC 209358) encoding human PRO207 (see AAY95338), which shows
XX CC homology to several members of the tumour necrosis factor family,
XX CC especially human lymphotoxin (23.4%). The cDNA was identified in a
XX CC foetal kidney cDNA library following identification of an expressed
XX CC sequence tag with homology to human Apo-2 ligand. A claimed method
XX CC for inhibiting the growth of a tumour cell comprises exposing the
XX CC tumor cell to PRO179, PRO207, PRO320, PRO219, PRO221, PRO224,
XX CC PRO328, PRO301, PRO356, PRO362, PRO356, PRO309 or PRO866 (see
XX CC AAY95337-49), their agonists or chimeric polypeptides incorporating
XX CC them. The tumour is especially a cancer selected from breast,
XX CC ovarian, renal, colorectal, uterine, prostate, lung, bladder and
XX CC central nervous system cancer, melanoma and leukaemia. Nucleic
XX CC acids encoding PRO179 etc. are used in the recombinant production
XX CC of the antitumour polypeptides.
XX SO Sequence 1353 BP; 257 A; 443 C; 389 G; 264 T; 0 other;
  
```

Alignment Scores: 3.96e-90 Length: 1353

QY 37 AAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeuLeu 56
 |||||||
 Db 61 GCCGCCCTCGAGAGCAGAGCGAGGGGCGCGGGGAGGCGCGGACCGCCCTGCTG 120
 QY 57 ValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaVal 76
 |||||||
 Db 121 GTCCCTCTCGGCTGGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGGCGCTGG 180
 QY 77 ValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluVal 96
 |||||||
 Db 181 GTGAGTTTGGGGAGCCGGGATCGCTCTCGCCAGAGAGCTTCCCAAGAGAGCTG 240
 QY 97 AlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAspPro 116
 |||||||
 Db 241 GCAGAGAGAGACAGACCGCGTGAATCCCAAGACAGAGAAACAGAGATCCT 300
 QY 117 AlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThr 136
 |||||||
 Db 301 GCCCTTTCTGAACCGACTGACTGCGCTCGCAGAAAGTCCACCTAAAGCGCGAAACA 360
 QY 137 ArgAlaArgArgAlaIleAlaIleAlaHisTyrGluValHisProArgProGlyGlnAspGly 156
 |||||||
 Db 361 CGGGCTCGAAGAGCGATCGCAGCCCATTTATGAGTTTCATCCAGAGACCTGGACAGAGGGA 420
 QY 157 AlaGlnAlaGlyValAlaAspGlyThrValSerGlyTyrPoliGluAlaArgIleAsnSerSer 176
 |||||||
 Db 421 GCCAGAGCAGAGTGGAGCGGACAGTAGTGGCTGGAGAGAGACCAAGATCAACAGCTCC 480
 QY 177 SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValHisThrArgAlaGlyLeuTyr 196
 |||||||
 Db 481 AGCCCTGCTGGCTACAAACCCCGAGATCGGGAGTTTATAGTTCACCCGGGCTGGCTCTAC 540
 QY 197 TyrLeuTyrCysGlnValHisPheAspGlyGlyValAlaValTyrLeuLysLeuAspLeu 216
 |||||||
 Db 541 TACCTGACTGATGAGGTGCACTTTGATGAGGAGGAGAGGCTCTCACTGAAGCTGGACTTG 600
 QY 217 LeuValAlaAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSer 236
 |||||||
 Db 601 CTGGTGAGATGCTGTCTGGGCGCTGCGCTGCTGGAGAAATTTTCAGACACTGGCGGAGT 660
 QY 237 SerLeuGlyProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuAlaLeuArgProGly 256
 |||||||
 Db 661 TCCCTGGGGCGCCAGCTCCGCTCTCCAGAGTGTGGCTGTGGCGCTGGCGGCGAGGG 720
 QY 257 SerSerLeuArgIleArgThrLeuProTyrPalaHisLeuLysAlaAlaProPheLeuThr 276
 |||||||
 Db 721 TCCCTCCCTGGGATCCGCGACCTCCCTCCGCGCCATCTCAAGCTGCGCCCTTCTCACC 780
 QY 277 TyrPheGlyLeuPheGlnValHis 284
 |||||||
 Db 781 TACTTCGAGACTTTCAGAGTTTCAC 804
 RESULT 6
 AAV47613
 ID AAV47613 standard; cDNA: 1236 BP.
 AC AAV47613;
 XX 27-OCT-1998 (first entry)
 DT
 XX TNF related endothelium proliferative agent gene.
 DE
 XX ss: TNF; endothelium proliferative agent; TREPA; wound healing; cancer;
 KM tissue grafting; vascularisation; apoptosis; autoimmune; birth control.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH 1..750
 CDS /tag= a
 FT /product= "TREPA"
 FT
 XX

PN W09835061-A2.
 XX 13-AUG-1998.
 PD
 XX 12-FEB-1998; 98WO-US02859.
 PF
 XX 10-FEB-1998; 98US-0021706.
 PR 12-FEB-1997; 97US-0798692.
 XX
 PA (ABBO) ABBOTT LAB.
 XX
 PI Wiley SR;
 XX
 DR WPI: 1998-447255/38.
 DR P-PSDB: AAM29745.
 XX
 PT Detecting nucleic acid encoding TREPA - useful for diagnosis and
 treatment of autoimmune disease, tumours and inflammation
 PS Claim 11; Page 123-4; 142pp; English.
 XX
 CC The TNF-related endothelium proliferative agent (TREPA), or its
 CC activators or agonists, are used to treat a deficit of TREPA, e.g. to
 CC promote wound healing or tissue grafting, by promoting vascularisation,
 CC also to induce apoptosis for treating cancer and eliminating autoreactive
 CC T cells, as an adjunct to cancer chemotherapy or antiviral treatment.
 CC TREPA peptides can also be used to target cytotoxic agents or for
 CC affinity isolation of the corresponding receptor, the nucleic acid for
 CC which can be used to transform tumour cells to render them more
 CC responsive to TREPA and to screen for TREPA mimics.
 CC Ribozymes, antisense RNA, antibodies or peptides, are used to treat
 CC TREPA-associated diseases, e.g. tumours and metastases (by inhibiting
 CC vascularisation), inflammation or a wide range of autoimmune conditions,
 CC conditions involving abnormal stimulation of epithelial cells (e.g.
 CC atherosclerosis), for birth control (inhibiting ovulation and placental
 CC formation) or other angiogenic conditions (e.g. ulcers).
 XX
 SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other;
 Alignment Scores:
 Pred. No.: 7, 37e-83 Length: 1236
 Score: 1265.00 Matches: 248
 Percent Similarity: 100.00% Conservative: 1
 Best Local Similarity: 99.60% Mismatches: 0
 Query Match: 87.60% Indels: 0
 DB: 19 Gaps: 0
 US-09-245-198a-4 (1-284) x AAV47613 (1-1236)
 QY 36 MetAlaAlaArgArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeu 55
 |||||||
 Db 1 ATGGCGCGCCGTCGAGAGCGAAGCGAGGGGCGCGGGGAGGCGCGGACCGCCCTG 60
 QY 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75
 |||||||
 Db 61 CTGGTCCCGCTCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGCTGGGCGC 120
 QY 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluVal 95
 |||||||
 Db 121 GTGCTAGTTTGGGGAGCCGGGATCGCTCTCGCCAGAGAGCTTCCCAAGAGAGCTG 180
 QY 96 ValAlaGlnGluAspGlnAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAsp 115
 |||||||
 Db 181 GTGGCAGAGAGAGACAGAGCCCGTGAATCCCAAGACAGAGAAACAGAGAT 240
 QY 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProGlyArgGlyThr 135
 |||||||
 Db 241 CCGTGGCTTTCTGAACCGACTGACTGCGCTCGCAGAAAGTCCACCTAAAGCGCGAAAA 300
 QY 136 ThrArgAlaArgArgAlaIleAlaIleAlaHisTyrGluValHisProArgProGlyGlnAsp 155
 |||||||
 Db 301 ACAGGGGCTCGAAGAGCGATCGCAGCCCATTTATGAGTTTCATCCAGAGACTGGACAGGAC 360

```
QY 156 GYAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPgluGluAlaArgIleAsnSer 175
DB 361 GGAGCGCAGGAGGTGTGACGGGACAGTGAAGTGGCTGGAGGAGCAAGATCAACAGC 420
QY 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195
DB 421 TCCAGCCCTTCGCGTACACCGCAGATCGGGAGTTATAGTCACCCGGGCTGGGCTC 480
QY 196 TyrTyrIleuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeu 215
DB 481 TACTACCTGTACTGTACAGTGCACCTTTATAGAGGGAGGAGGCTGTCTACCTGAAGCTGGAC 540
QY 216 LeuLeuValAspGlyValIleuAlaLeuArgCysLeuGluGluPheSerAlaThrAla 235
DB 541 TTGCTGGTGGATGGTGTCTGGCCCTGGCTGGCTGGAGGAATCTCAAGCCATGCGGCG 600
QY 236 SerSerLeuGlyProGlnIleuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgPro 255
DB 601 AGTTCCTCGGAGCCCGACCTCGCTGCGCAGAGTGTGGCTGTGGCCCTGGGCGCA 660
QY 256 GlySerSerLeuArgIleArgThrLeuProTrrPalaHisLeuLysAlaAlaProPheLeu 275
DB 661 GGGTCTCTCCGTCGGATCCGACACCTCCCTGGGCCATCTCAAGGCTGCCCTTCTC 720
QY 276 ThrTyrPheGlyLeuPheGlnValHis 284
DB 721 ACCTACTCTCGGACCTCTCCAGGTTCCAC 747

RESULT 7
AAd04350
ID AAd04350 standard; cDNA; 1236 BP.
AC AAd04350;
XX 04-JUN-2001 (first entry)
DE Human TREPA (TNF related endothelium proliferative agent) cDNA.
XX Human: tumour necrosis factor; TNF; angiogenesis; wound healing;
KW TREPA: TNF related endothelium proliferative agent; tumour; metastasis;
KW grafting; vulnetary; ss.
XX Homo sapiens.
OS
FH Key 1..750 Location/Qualifiers
FT CDS /tag= a
FT product= "Human TREPA (TNF related endothelium
FT proliferative agent)"
PN US6207642-B1.
XX 27-MAR-2001.
PD
XX 26-JUN-1998; 98US-0105343.
XX 12-FEB-1997; 97US-0798692.
PR 10-FEB-1998; 98US-0021706.
XX (ABBO ) ABBOTT LAB.
PA
XX WAlley SR;
PI
XX WPI: 2001-280760/29.
DR P-PSDB: AAE00891.
XX Inducing angiogenesis in mammal at desired sites for promoting wound
XX healing, by administering soluble fragment of extracellular domain of
XX tumor necrosis factor related endothelium proliferative agent protein
XX Example 2; Column 73-74; 53pp; English.
```

```
CC The present invention relates to extracellular signal molecules,
CC particularly members of tumour necrosis factor (TNF) family molecules
CC designated as TREPA (TNF related endothelium proliferative agent),
CC Soluble biologically active TREPA are used to treat TREPA-associated
CC diseases, tumours or metastases. TREPA is used for inducing angiogenesis
CC in human for promoting wound healing and for vascularising grafted
CC tissue for successful grafting and to promote tissue grafts.
CC The present sequence is a cDNA clone ID #690050 encoding human TREPA.
XX
SQ Sequence 1236 BP; 225 A; 416 C; 358 G; 237 T; 0 other:

Alignment Scores:
Pred. No.: 7.37e-83 Length: 1236
Score: 1265.00 Matches: 248
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.60% Mismatches: 0
Query Match: 87.60% Indels: 0
DB: Gaps: 0

US-09-245-198A-4 (1-284) x AAd04350 (1-1236)
QY 36 MetaAlaAlaArgSerGlnArgArgGlyArgArgGlyGluProGlyThrAlaLeu 55
DB 1 ATGGCCGCCCGTCGGAGCCAGAACGGAGGGGCGCCGGGGGAGACCCGGCCCTG 60
QY 56 LeuValProLeuAlaLeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAla 75
DB 61 CTGGTCCCGCTCGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 120
QY 76 ValValSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeu 95
DB 121 GTGGTCAGTTGGGAGCGCGGATCGCTGTCGCCCGCCAGAGCCTGCCAGAGAGCTG 180
QY 96 ValAlaGluGluAspGluAspProSerGluLeuAsnProGlnThrGlnGluSerGlnAsp 115
DB 181 GTGGCAGAGAGGAGCAGACCGCGTGGAGACTGAATCCCGACAGCAAGAAAGCCAGAT 240
QY 116 ProAlaProPheLeuAsnArgLeuValArgProArgArgSerAlaProLysGlyArgLys 135
DB 241 CCTGGCGCTTTCCTTAACAGCACTAGTTGGCTCGAAGAGCCACTTAAGGCCGGA 300
QY 136 ThrArgAlaArgArgAlaIleAlaHisTyrGluValHisProArgProGlyGlnAsp 155
DB 301 ACACGGGCTCGAAGAGCATGCGACCCCATATGATTCATCCAGCATGAGACAGGAC 360
QY 156 GYAlaGlnAlaGlyValAspGlyThrValSerGlyTyrPgluGluAlaArgIleAsnSer 175
DB 361 GGAGCGCAGGAGGTGTGACGGGACAGTGAAGTGGCTGGAGGAGCAAGATCAACAGC 420
QY 176 SerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeu 195
DB 421 TCCAGCCCTTCGCGTACACCGCAGATCGGGAGTTATAGTCACCCGGGCTGGGCTC 480
QY 196 TyrTyrIleuTyrCysGlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeu 215
DB 481 TACTACCTGTACTGTACAGTGCACCTTTATAGAGGGAGGAGGCTGTCTACCTGAAGCTGGAC 540
QY 216 LeuLeuValAspGlyValIleuAlaLeuArgCysLeuGluGluPheSerAlaThrAla 235
DB 541 TTGCTGGTGGATGGTGTCTGGCCCTGGCTGGCTGGAGGAATCTCAAGCCATGCGGCG 600
QY 236 SerSerLeuGlyProGlnIleuArgLeuGlyGlnValSerGlyLeuLeuAlaLeuArgPro 255
DB 601 AGTTCCTCGGAGCCCGACCTCGCTGCGCAGAGTGTGGCTGTGGCCCTGGGCGCA 660
QY 256 GlySerSerLeuArgIleArgThrLeuProTrrPalaHisLeuLysAlaAlaProPheLeu 275
DB 661 GGGTCTCTCCGTCGGATCCGACACCTCCCTGGGCCATCTCAAGGCTGCCCTTCTC 720
QY 276 ThrTyrPheGlyLeuPheGlnValHis 284
DB 721 ACCTACTCTCGGACCTCTCCAGGTTCCAC 747
```


DR WPI: 1998-145619/13.
 DR P-PSDB; AAW47524.
 XX Tumour necrosis factor related ligand - useful for, e.g. treating
 PT cancer, auto-immune disease and immune responses to tissue grafts
 PS Claim 2; Pages 45-46; 69pp; English.
 XX The sequence is that encoding mouse tumour necrosis factor related
 CC ligand (TRELL). TRELL or active fragments can be included with a
 CC carrier in pharmaceutical compositions to treat cancer, autoimmune
 CC diseases or immune responses to tissue grafts, or to stimulate or
 CC suppress the immune system. It is useful to screen for TRELL
 CC receptors, by labelling with a detectable label and screening
 CC compositions for binding. Agents interfering with TRELL-receptor
 CC binding can also be screened for, can then be administered,
 CC optionally with interferon- gamma, to induce cell death or
 CC treat, suppress or alter immune responses (especially involving human
 CC adenocarcinoma cells) involving a signal pathway between TRELL and its
 CC receptor. The DNA sequence can be used in gene therapy for
 CC TRELL-related disorders in mammals (especially humans), e.g. tumours,
 CC autoimmune and inflammatory diseases or inherited genetic disorders,
 CC by introducing into cells, and expressing, therapeutically effective
 CC amounts of a vector, e.g. a virus comprising a gene encoding TRELL.
 CC It may also be of use in the preparation of probe probes for
 CC screening natural/synthetic DNAs for TRELL-encoding sequences
 CC and for antisense therapy.
 XX
 XX Sequence 1168 BP; 242 A; 360 C; 298 G; 268 T; 0 other;
 Alignment Scores:
 Pred. No.: 3.8e-65 Length: 1168
 Score: 1020.00 Matches: 199
 Percent Similarity: 92.86% Conservative: 9
 Best Local Similarity: 88.84% Mismatches: 16
 Query Match: 70.64% Indels: 0
 Gaps: 0
 DB: 19
 US-09-245-198a-4 (1-284) x AAW18599 (1-1168)
 QY 61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuAlaValaValaSerLeuGly 80
 DB 5 CTAGAGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGGCTGGGG 64
 QY 81 SerArgLaserLeuSerArgLaserArgLaserArgLaserArgLaserArgLaserArg 100
 DB 65 AGCTGGGCAAGCTGCTGCTGGGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 124
 QY 101 GlnAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeu 120
 DB 125 CGGAGAGCCCTGAACTGAATCCCGACAGAGAAAGCCAGATGTGTCTTCTTG 184
 QY 121 AsnArgLeuValArgProArgSerArgProArgSerArgSerArgSerArgSerArg 140
 DB 185 GAAACAATACCTCGGGCTCGAAGAAGTCTCTTAAGCCGGGAAGGGCGGCTCGCGCA 244
 QY 141 AlaIleAlaIleHisIleGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGly 160
 DB 245 GCATATTCAGAGCCCATATGAGGTTCATCTCGGCGAGAGAGATGAGAGCAAGCAGGT 304
 QY 161 ValAspGlyThrValSerGlyTTPGluGluAlaAlaArgIleAsnSerSerSerProLeuArg 180
 DB 305 GTGAGTGAGGACAGTGAAGTGTCTGGGAGAGCAACCAATTCACAGCTCCAGCTCTCGCG 364
 QY 181 TyrAsnArgIleIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200
 DB 365 TACGACCGCCAGATGGGGAATTTACAGTCATGAGGGCTGGGCTGACTACTGTACTGT 424
 QY 201 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220
 DB 425 CAGGTGACATTGTATGAGGAAAGCGCTGTACTCTGAAGCTGAGCTTGTGCTGTAACGGT 484
 QY 221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSerSerLeuGlyPro 240

DB 485 GTCTGGCCCTCGGCTGCGGAATTCAGCCACAGCAGACAGCTCTCGGGCCC 544
 QY 241 GlnLeuArgLeuGlyGlnValSerGlyLeuAlaLeuArgProGlySerSerLeuArg 260
 DB 545 CAGCTCCGTTTGGCCAGGAGGTCTGTGCGGCGGCGGAGGAGGTCTTCCCTTCGG 604
 QY 261 IleArgThrLeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeu 280
 DB 605 ATCGGACCCCTCCCGTGGCTCATCTTAAGGTGCGCCCTTCTTAACCTTCTTGACCTC 664
 QY 281 PheGlnValHis 284
 DB 665 TTTCAGACTCAC 676
 RESULT 11
 ID AAX23425 standard; DNA: 701 BP.
 AC AAX23425;
 DT 18-JUN-1999 (first entry)
 XX Mouse TNRL3 DNA.
 XX Tumour necrosis factor receptor; signal transducer molecule; TNF; APO4;
 KW developmental abnormality; gestational abnormality; prostate cancer;
 KW APO6; APO8; APO9; TNRL-1; TNRL-3; diagnosis; treatment; therapy; disease;
 KW cytoplasmic domain; immunogen; antibody preparation; breast carcinoma;
 KW apoptosis; mouse; ss.
 XX Mus sp.
 OS
 XX Key Location/Qualifiers
 FH 1.636
 FT CDS /tag= a
 FT /product= "TNRL3"
 FT
 XX WC0911791-A2.
 XX
 PD 11-MAR-1999.
 PE 04-SEP-1998; 98MO-US18393.
 XX
 XX 05-SBP-1997; 97OS-0924634.
 XX
 PA (UNITW) UNIV WASHINGTON.
 XX
 PI Chaudhary PW;
 XX
 DR WPI: 1999-205191/17.
 DR P-PSDB; AAW93591.
 XX
 PT New Tumor Necrosis Factor family receptor polypeptides and ligands -
 PT useful for diagnosis and treatment of prostate cancer and
 PT developmental or gestational abnormalities
 XX
 XX Example VII; Fig 13B; 156pp; English.
 PS
 CC This invention describes isolated Tumor Necrosis Factor (TNF) family
 CC receptor polypeptides: APO4, APO6, APO8 and APO9 or their active
 CC fragments, and isolated TNF related ligands 1 and 3 (TNRL1 and TNRL3) or
 CC their active fragments. APO4 is useful for diagnosing prostate cancer
 CC by determining levels of APO4 in an individual. Prostate cancer can also
 CC be treated using APO4 selective binding agents linked to a therapeutic
 CC moiety. APO4 polypeptides are also useful for identifying selective
 CC binding agents, useful in diagnosis/treatment of disease by binding of
 CC agents to the polypeptide/active fragment which is extracellular, or
 CC expressed on the cell surface. The binding is preferably performed in
 CC vivo. APO4 polypeptides/active fragments are also useful for screening
 CC for agonists and antagonists by binding and observing the change in APO4
 CC activity. Effective pharmacological agents useful in diagnosis or
 CC treatment of disease are also identified using APO4 polypeptides/active

fragments and APO4 signal transducer molecules that specifically interact with a cytoplasmic domain of APO4 and detecting a change in level of APO4 activity. The method is performed in vivo or in vitro. APO polypeptides are all useful as immunogens for preparing antibodies. APO4 is also useful for diagnosis/treatment of developmental or gestational abnormalities. APO8 was transfected to human breast carcinoma cell line MCF-7, and induced apoptosis.

SO Sequence 701 BP; 139 A; 210 C; 203 G; 149 T; 0 other;

Alignment Scores:

Pred. No.:	1,25e-61	Length:	701
Score:	968.00	Matches:	188
Percent Similarity:	93.36%	Conservative:	9
Best Local Similarity:	89.10%	Mismatches:	14
Query Match:	67.04%	Indels:	0
DB:	20	Gaps:	0

US-09-245-198A-4 (1-284) x AAV23425 (1-701)

```

OY 74 LeuAlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGlu
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 CTGCTCGTGTGAGCGTGGAGCGGAGCGTGTCTGCCAGACCTTCTCAGAG 60

OY 94 GluLeuValAlaGluGluAspGlnAspProSerGluLeuAsnProGlnThrGluGluSer
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GAGCTGACAGCAGAGAGCCGCCGAGCCCTGACTGAAATCCCCACAGAGAAAGC 120

OY 114 GlnAspProAlaProPheLeuAsnArgLeuValArgProArgSerAlaProGlySly
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 CAGGATGTGATCTTCTTGGAACTAGTCCGCGCTCGAAGAAAGTCTCTAAAGC 180

OY 134 ArgLysThrArgAlaArgArgAlaAlaAlaHisTyrGluValHisProArgProGly
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 CGGAAGCGCGCCCTCGCGGAGCTATTGACCCCTTTGAGGTTCATCTCGGACAGA 240

OY 154 GlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTyrGluGluAlaArgIle
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 CAGGTGAGCAGCAAGCAGGTGTGATGGACAGTACTGCTGGAAAGACCAAAATC 300

OY 174 AsnSerSerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAla
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 AACAGCTCACCCCTCTGCTACAGCCGCCAGATTGGGGAATTACATCAGAGGCT 360

OY 194 GlyLeuTyrTyrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTyrLeuIys
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GGGCTTACTACTGTACTGTGAGGTGCACTTGTATGAGGAAAGGCTGTCTACTGAG 420

OY 214 LeuAspLeuValAspGlyValLeuAlaLeuArgCysLeuGluGluPheSerAlaThr
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 CTGGACTTGTGCTGTAACAGGTGTGTGCTGCCCTGCCCTGGAAATTTCTCAGCCACA 480

OY 234 AlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGlnAlaSerGlyLeuLeuAlaLeu
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 481 GCACCAACCTCTCTGCGGCCACAGCTCGTTGTGCCAGTGTCTGGGCTGTGGCGCTG 540

OY 254 ArgProGlySerSerLeuArgIleArgThrLeuProTyrAlaHisLeuAlaAlaPro
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 541 CGGCCAGGGTCTCCCTTCGGATCCGACCCCTCCCTGGCTCATTAAGCTGCCGCC 600

OY 274 PheLeuThrTyrPheGlyLeuPheGlnValHis 284
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 601 TTCTTAACCTTGTGACCTTTTCAAGTTCAAC 633

RESULT 12
AAV24140
ID AAV24140 standard; DNA; 2856 BP.
XX
AC AAV24140:
XX
DT 28-SEP-1998 (first entry)
XX
DE Homo sapiens BARD1 exon 1.

```

```

XX KW BARD1: ring protein; BRCA1: breast cancer; risk; diagnosis; exon; ss.
XX OS Homo sapiens.
XX FH Key
XX FT exon
XX FT /tag= a
XX FT /number= 1
XX PN W09812327-A2.
XX PD 26-MAR-1998.
XX PF 19-SEP-1997; 97WO-US16842.
XX PR 04-APR-1997; 97US-0042985.
XX PR 20-SEP-1996; 96US-0025296.
XX PR 03-APR-1997; 97US-0042611.
XX PA (TEXA ) UNIT TEXAS SYSTEM.
XX PI Baer R, Bowcock AM;
XX DR WPI: 1998-230317/20.
XX PT DNA sequence encoding BARD1, B123, BE2, BE14, BE31 or BE445 - which
XX PT as breast cancer antigen, BRCA1, binding proteins are useful to
XX PT identify patient having or at risk of developing cancer
XX PS Claim 25; Page 307-308; 348pp; English.
XX CC The sequence is that of containing a BARD1 exon which can be used in the
XX CC preparation of the recombinant breast cancer antigen, BRCA1, binding
XX CC proteins BARD1, B123, BE2, BE14, BE31 or BE445, or a composition for the
XX CC detection of a BARD1, B123, BE2, BE14, BE31 or BE445 nucleic acid
XX CC sequence, specifically a wild type BARD1 composition for the detection
XX CC or purification of BRCA1, useful to identify a patient having, or at
XX CC risk of developing cancer. BARD1 can be used in the preparation of an
XX CC anti-BARD1 antibody, and in the detection and purification of a BRCA1
XX CC protein. BARD1, B123, BE2, BE14, BE31 or BE445 can be used in the
XX CC identification of a binding protein agonist or antagonist that alters
XX CC the binding of BARD1, B123, BE2, BE14, BE31 or BE445 to BRCA1 or the
XX CC biological activity of the BRCA1-BARD1, B123, BE2, BE14, BE31 or BE445
XX CC complex. The antibodies can be used to detect BARD1, B123, BE2, BE14,
XX CC BE31 or BE445, a specific anti-BARD1 antibody can be used to identify
XX CC a patient having or at risk of developing cancer.
XX SO Sequence 2856 BP; 616 A; 778 C; 786 G; 674 T; 2 other;

Alignment Scores:
Pred. No.: 4.74 Length: 2856
Score: 127.00 Matches: 84
Percent Similarity: 36.00% Conservative: 33
Best Local Similarity: 25.85% Mismatches: 122
Query Match: 8.80% Indels: 86
DB: 19 Gaps: 13

US-09-245-198A-4 (1-284) x AAV24140 (1-2856)
OY 14 ProLeuProArgSerLeuGlySerArgAspGlyGlyAlaVal----- 27
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1788 CCCACTGCGCCAGCGGCTGTGCGGCCGCCGCCCTGTGCTCGGAGTCCCTAATT 1847

OY 28 -----ArgGlnAlaGlnPro-----ProAlaProMetaAla 38
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1848 TGGAGCATTCGCGCGCGCTGCGGCCGCCGCCGCCGCCGCCCTGTGCGGCC 1907

OY 39 ArgArgSerGlnArgArgArgArgGlyArgGlyGluProGlyThrAlaLeuLeuValPro 58
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1908 CGCGTCCCAAGACGCGGAGAGCTTGGCGGTTCGAGTGCCTGACGCTCTCCCT 1967

OY 59 LeuAlaLeuGlyLeuAlaLeuAlaLeuAlaLeuAlaLeuAlaValSer 78

```

```

Db 1968 -----
OY 79 LeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluLeuValAlaGlu 98
Db 1977 CGAGGCGCTCTTGGCTCCGCTTCCGAGAGAGCTTTCATCCGAAGCGGAGCATGCCG 2036
OY 99 GluAspGlnAspProSerGlnLeuAsnProGlnThrGlnGluSerGlnAspPro----- 116
Db 2037 GATTAATCGGACAGCGAGAACCGGACCGGAGATCCGCTCCGCGAAGACGCTCGTTCC 2096
OY 117 AlaProPheLeu-AsnArgLeuValArgProArgArgSerAlaProGlyArgLysThr 136
Db 2097 GCGCGCCGATGAGAACCGGATGTCGCGGTCGCGGCCACAGTCGCGCGCGCTCGAG 2156
OY 136 ArgAlaArgArgAlaAlaAlaHisThrGluValHisProArgProGlyGlnAspG 156
Db 2157 CCGCTGGAGAGAGCTGCTGGCTCT-----CGCGTTGGTAAAGACGG 2198
OY 156 ValAlaGlnAlaGly-----ValAspGlyThrValSerGlyTTPGlnGluAlaArg 173
Db 2199 AGCTTCTGGGGTGGCTCGAGCGAGCGGTCGACAGTTTCTGGGGCGCAGATCT 2258
OY 173 eAnSerSerSerProLeuArgTyrAsnArgGlnIle----- 185
Db 2259 TTTCAATCTCTCGTTCTCTCCGCTCCGCGCTGCAATCGGCTCGCGCTTA 2318
OY 186 -----GlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGln 202
Db 2319 GCACCTCGCGGGGATATAGTATTAACACTTCTGCTTCATCTCACTTATTTTGG-- 2376
OY 202 HisPheAspGlyGlyLysAlaValTyr-----LeuLysLeuAspLeuValAs 219
Db 2377 -----GCGACTTACCGCGCTCCCTTCCCTTCCCTGATCCCACTGAAC 2417
OY 219 pGlyValIleLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaSerSerLeu-G 239
Db 2418 GGTAGTTTTTGAACCTTCACGCGGCTGAGAACCGTCTGAGAGTGCTGAATAAAATGTT 2477
OY 239 1yProGlnLeuArgLeuCysGlnValSer-----GlyLeuLeuAlaL 253
Db 2478 CATCCCGGCTCGCGCTCCAGAGTTTGAATCGGCTGGGGGCTGAGGCTTCTGCATT 2537
OY 253 euArgProGlySerSerLeuArgIleArgThrLeuPro-----TrpAla- 267
Db 2538 TTTTACCGCGCTCGATTATACC--CGCTGCTTCCGGAGCTGAGGCAATTTGGCTG 2594
OY 268 -----HisLeuLysAlaAlaProPheLeuThrTyrPheGlyL 280
Db 2595 GCGGGCGCGCGCGAGACCTTAATTTAGAACGACTGCCACTTAAGTTAACTGGCC 2654
OY 280 euPheGlnVal 283
Db 2655 TTTTGTGACATT 2665

RESULT 13
ID AAA49205
ID AAA49205 standard; cDNA; 1660 BP.
AC AAA49205;
XX
XX 12-DEC-2000 (first entry)
DE Corn putative lecithin:cholesterol acyltransferase gene #5.
XX
XX Corn; lecithin:cholesterol acyltransferase; phytosterol;
KM phosphatidylcholine-sterol O-acyltransferase; heat shock; cold shock; ss.
XX
XX zea mays.
XX
XX Key Location/Qualifiers
FH 77.1396
CDS /tag= a
FT

```

```

FT
FT /product="lecithin:cholesterol acyltransferase"
FT /EC_number="2.3.1.43"
PN MO200032791-A2.
XX
XX 08-JUN-2000.
PD
XX
XX 02-DEC-1999; 99WO-US28586.
PE
XX
XX 03-DEC-1998; 98US-0110782.
PR
XX
XX (DUPO ) DU PONT DE NEMOURS & CO E. I.
PA
PA Cahoon RE, Kinney AJ, Sakai H, Shen JB, Butler KH, Saylor JT;
PI
PI WPI; 2000-412337/35.
DR P-PSDB; AAB01210.
XX
XX PT Polynucleotide encoding plant lecithin:cholesterol acyltransferase
XX PT enzyme useful for producing transgenic plants and for producing
XX PT antibodies specific to which is useful for screening cDNA expression
XX PT libraries
PS Claim 2; Page 43-44; 49pp; English.
XX
XX The present sequence is a putative coding sequence for a corn
XX lecithin:cholesterol acyltransferase (also known as
XX phosphatidylcholine-sterol O-acyltransferase). This enzyme is found
XX associated with high-density lipoproteins and Apolipoprotein-AI and -D.
XX The gene and protein can be used to produce transgenic plants which have
XX increased lipid metabolism and membrane fluidity, and therefore increased
XX resistance to heat and/or cold shock, to alter the content of phytosterol
XX or lecithin in grains and to identify potential herbicides.
SQ Sequence 1660 BP; 361 A; 475 C; 499 G; 325 T; 0 other;

Alignment Scores:
Pred. No.: 2.81 Length: 1660
Score: 126.50 Matches: 81
Percent Similarity: 34.48% Conservative: 29
Best Local Similarity: 25.39% Mismatches: 109
Query Match: 8.76% Indels: 100
DB: Gaps: 16

US-09-245-198a-4 (1-284) x AAA49205 (1-1660)
OY 11 ArgArgLeuProLeuProArgSerLeuGlySerArgSpGlyValAlaVal----- 27
Db 394 CAACGAGTACCGGACCTGCGCG-----CGTCGAGACGCGAGTCCCAACTTGGCTC 447
OY 28 --ArgGlnAlaGlnProProAlaPro-----MetaAlaAlaArgArgSerGln 42
Db 448 CACAGAGGCTTCACGACAGAACCCCTCAAGTCAGACTGGCTGCGGAAGCTGAG 507
OY 43 ArgArgArgGlyArgArgGlyGluProGlyThrAlaLeuLeuVal-----ProLeuAla 60
Db 508 AGCCGCACTGAGAGACATGGATCCGAGACGAGACACCATTTTGGAGCCCTACGA 567
OY 61 LeuGlyLeuGlyLeuAlaLeu-----AlaCysLeuGlyLeuLeuLeu 74
Db 568 CTTCGCTAGCGCGCGCGCTCCCGCGCCAGACAGCTCCGAGGTACTCCCGCTACTTCAA 627
OY 75 AlaValAlaSerLeuGlySerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGlu 94
Db 628 GAGCGTGAATCGACACTGCTGCA----- 648
OY 95 LeuValAlaGlnGluGlnAspGlnAspProSerGlnLeuAsnProGlnThrGlnGluSerGln 114
Db 649 -----GCGCGGAGCGAGGAGGACCG-----GAGGAAGCGCGT 681
OY 115 AspProAlaProPheLeuAsnArg-----LeuValArgProArgArgSerAla 130
Db 682 CATCTCGGCGCACAGCTTGGCGGATGTGCTGCGAGTTCTCGGAAACACTCGGCC 741

```

```
OY 131 ProlYsGIARgLYsThrARgAlARgAlAlAlAlAlAlAlAlAlAlHisPro 150
DB 742 -----GGCGTGGCGCGCGGACACATCGACG-----CCT 771
OY 151 ArgProGIyGln-----AspGIyAlAGln----- 158
DB 772 CGTCCTGGTGGCGCGCGCGGCTCCGAGCGCGGCGGACATTGGCGTC 831
OY 159 -----AlaGIyAlAspGIyThrValSer 166
DB 832 CGGAGCGACATCTCTACGTCCAGCAGCAGCGCGCTGGCCAGCGCATGTGGAG 891
OY 167 -GIYTPGIuGIuAlaArgIleAsnSerSerSerProLeuAlaArgIleGlnIleG1 186
DB 882 GAGCTTCGAGAGCGGACATCGTAACCTCCGCGCGCGCGCTGGCGGCGCTGCAGGC 951
OY 186 yGIuPheIleValThrArgAlaGIyLeuTYrTYrLeuTYrCysGlnValHisPheAspG1 206
DB 952 GCCCGCTCGTGGTACACAGGAGCGGACACTAC----- 982
OY 206 uGIyLysAlaValTYrLeuLysLeuAspLeuValAspGIyValLeuAlaLeuArgCy 226
DB 983 ----TCCCGCTCCGCGCAGCAGCATGGAGCGCTCTCCGCGCGCGCTGGCTCCGCGGAGGC 1038
OY 226 sLeuGIuGIuPheSerAlaThrAlaAlaSerSerLeuGIy-----ProGlnLe 242
DB 1039 CGCGAGCCCTTCAGAGACAGCGGCGCTCCCAAGATGGCGACTTCGCGCGCGCATGTGT 1098
OY 242 uArgLeuCysGlnValSerGIyLeuLeuAlaLeuArgProGIySerSerLeuArgIleAr 262
DB 1099 GCCCATGACGTACATCAGCGGGGTC-----GGCAACAGAGACCGCGCTGGC 1143
OY 262 gThreuProTYrAlaHis-----LeuLysAlaAlaProPheLeuThrTYrPheGIy 279
DB 1144 GCTGGTGTGTGGGCGAAGACTTCGACGCGGCGCGGAGGTGGCGCGCTACGG 1198

RESULT 14
AAV69289
ID AAV69289 standard; cDNA: 2832 BP.
XX
AC AAV69289;
XX
DT 01-MAR-1999 (first entry)
XX
DE Murine JNK-interacting protein 1 (JIP-1) cDNA.
XX
KW JIP-1; JNK-interacting protein; c-Jun NH2-terminal kinase;
KW signal transduction; inhibitor; mouse; neurodegenerative disease;
KW Parkinson's disease; Alzheimer's disease; blood clot; stroke;
KW malignancy; cancer; leukaemia; autoimmune disease; inflammation;
KW apoptosis; therapy; diagnosis; ds.
XX
OS Mus sp.
XX
FH Key Location/Qualifiers
FT CDS 180..2162
FT /tag= a
XX
XX MO9849188-A1.
XX
XX 05-NOV-1998.
XX
XX 28-APR-1998; 98WO-US08513.
XX
XX 28-APR-1997; 97US-0819177.
XX
XX (UYMA-) UNIV MASSACHUSETTS.
XX
XX Davis RJ, Dickens M;
XX
XX WPI: 1999-024042/02.
XX
XX P-PSDB: AAW81525.
```

```
XX
PT c-Jun NH2-terminal kinase (JNK)-interacting protein 1 - used to
PT treat neurodegenerative disease, blood clot, leukaemia, autoimmune
PT disease, and inflammation
XX
XX Claim 11; Page 66-70; 95pp; English.
XX
CC This cDNA sequence codes for murine JNK-interacting protein 1
CC (JIP-1; see AAW81625), a novel cytoplasmic anchor protein that
CC specifically binds to and inhibits the biological effects of JNK
CC (c-Jun NH2-terminal kinase), including the initiation of apoptosis
CC and oncogenic transformation. To isolate the clone, a yeast
CC two-hybrid method was used to screen a mouse embryo cDNA library to
CC identify proteins that interact with JNK. 7 overlapping partial
CC clones were isolated, and the full-length cDNA was subsequently
CC obtained by screening a mouse brain lambda ZAP11 cDNA library. The
CC invention provides JIP-1 nucleic acids and polypeptides, expression
CC vectors and host cells. The JIP-1 polypeptide and nucleic acids
CC (including antisense and ribozyme molecules) can be used in the
CC manufacture of a medicament for treating a pathological condition
CC associated with abnormal expression or activity of JNK, such as a
CC neurodegenerative disease (selected from Parkinson's disease and
CC Alzheimer's disease), a blood clot, stroke, malignancy, leukaemia,
CC an autoimmune disease or inflammation (all claimed).
XX
SQ Sequence 2832 BP; 569 A; 888 C; 840 G; 535 T; 0 other:
XX
XX
XX Alignment Scores:
XX Pred. No.: 7.12 Length: 2832
XX Score: 124.50 Matches: 72
XX Percent Similarity: 36.06% Conservative: 25
XX Best Local Similarity: 26.77% Mismatches: 73
XX Query Match: 8.62% Indels: 100
XX DB: Gaps: 12
XX
US-09-245-198a-4 (1-284) x AAV69289 (1-2832)
OY 1 MetSerIleuAspPheGluIleSerAlaArgLeuProLeuProArgSerLeuGIy 20
DB 20 ATAAAGCTTATATTCGCTGTCGCGAGCGCG-----CCGCGCCAGCTAGTCCGA 70
OY 21 SerArgAspGIyAlaValArgGlnAla----- 30
DB 71 CCCCGGGGCGGCGCGGCTCTCCACAGCGGCTCGCGCGCGCGCGCGCGCG 130
OY 31 GlnProProAlaProMetAlaAlaArgSerGln-----ArgArgArgGIyArg 47
DB 131 CGCCTCCCGCGCGCGCTCTCCCGGATGGCCAGGCTGCACCCCGAATGGCGGAGCG 190
OY 48 ArgGIyGIuProGIyThrAlaLeuLeuValProLeu-Ala-----LeuGIyLe 63
DB 191 AAGAGACGGCTGGGGCGGCGCGCGCTCCACCGCGCGCTCCCATTCCTGGGACT 250
OY 63 u----- 63
DB 251 GCACATCGCGTGGCTCCCAATTTCAGGCTCACCCATGACATCAGCTGAGAGATTGA 310
OY 64 -----GlyLeuAlaLeuAlaCys----- 69
DB 311 GGATGAAGACCTTTCGGAGATCAGTACAGAGTGTGCATCAGCTGCAAGACAC 370
OY 70 -LeuGIyLeu-----LeuLeuAlaValSerLeuGIySerArgAl 83
DB 371 CCTGTCTTCGCGCGCGCGCGCGGCTGCTGTCTGGGGGTAGCAGCGGACCGCGG 430.
OY 83 aSerLeuSerAlaGlnGIuProAlaGlnGIuLeuVal----- 96
DB 431 GAGCGCGCTGACGCGGAGATGCTGCAGATGACCTGTCGACGCGGAGTGCACATCC 490
OY 97 ----AlaGIuGIuAspGlnAsp-----ProSerGI 105
DB 491 GGGCGCGGAGACGAGAGAGAGAGAGACGACGAGCTCGTCCCAACGACGAGAGT 550
```

OY	105	uenuhsnp:progIntnThGlnGluSerGlnAspProAlaProPheLeuhsnMrGluValAr	125
Db	551	GGGGCCCTCCCAAGCGGAGTCCACACAGATCCGGCCCT-----	590
OY	125	pGProARgArSerAlaProLysGlyArguYThrArgAlaArgARgAlaIleAlaAlA1	145
Db	591	-----CGCAGCAGAGGGCGGGCC-GGCGACAGGCACGGAGACACCTACCGACCA	642
OY	145	sTyrgtLuVaIHsPProARgProGlyInAspGlyAlaGlnAlaGlyAlaAspGly-Thrv	165
Db	643	AGAGGCGTACACGCGTCAACCTTTTCCCGAGGT-----GCCGCGGTCTCAGACACAGC	696
OY	165	AsIser-----GlyTrpGluGluAlaArgIleAsnSerSers	177
Db	697	TCAAATATATACTCTTTTAGCCAAAGCACAGTGGCAGACCGGTGTCTCGATCATCT	756
OY	177	exProLeuARgTYrAsnARgIn	184
Db	757	CCCCTGTGAAGACAGGAGAACAG	779
RESULT 15			
AAD25519/C			
ID	AAD25519	standard; DNA; 154746 BP.	
XX	AAD25519;		
AC			
XX	26-MAR-2002	(first entry)	
DT			
XX			
DE		Human herpesvirus 2 complete DNA genome.	
XX			
XX			
KM		Human herpesvirus 2; cytostatic; cancer; immunosuppressive; virucide;	
KM		antibacterial; fungicide; protozoicide; antihelmatic; antiinflammatory;	
KM		antiarthritic; rheumatoid arthritis; neuroprotective; multiple sclerosis;	
KM		immune response; vasotropic; vaccine; gene therapy; autoimmune disease;	
XX		vasculitis; ds.	
XX			
OS		Human herpesvirus 2.	
XX			
PN		WO200176643-A1.	
XX			
PD		18-OCT-2001.	
XX			
PF		06-APR-2001; 2001WO-US11372.	
XX			
PR		07-APR-2000; 2000US-195680P.	
XX			
PA		(BAYU) BAYLOR COLLEGE MEDICINE.	
PI		Orson FM, Kinsey BM, Bhogal BS;	
XX			
DR		WPI: 2002-066308/09.	
XX			
PT		Composition for oral delivery of vaccines, comprises expression vector	
PT		containing antigenic genomic sequence, bound to aggregated	
PT		protein-polycationic polymer conjugate or suspension	
XX			
PS		Disclosure: Page 90-132; 145pp; English.	
XX			
CC		The invention relates to a composition comprising an expression vector	
CC		bound to an aggregated protein-polycationic polymer conjugate or	
CC		suspension. The expression vector contains a promoter polynucleotide	
CC		sequence operatively linked to a polynucleotide sequence encoding an	
CC		antigen which is a fragment of a gene or genome associated with an	
CC		infectious disease, cancer and autoimmune disease such as rheumatoid	
CC		arthritis, vasculitis, and multiple sclerosis, pathogenic genomes	
CC		consisting of bacterium, fungus, protozoa and virus such as human	
CC		immunodeficiency virus (HIV), herpes simplex virus (HSV), hepatitis C	
CC		virus (HCV), influenza and respiratory syncytial virus (RSV), and	
CC		optionally comprising a nucleotide sequence encoding a cytokine (or a	
CC		cyclokin expression vector). is useful for inducing an immune response	
CC		(systemic and/or mucosal) in an organism. The cytokine expression vector	
CC		contains a sequence for granulocyte macrophage colony stimulating factor	
CC		(GM-CSF) or interleukin-12 (IL-12). The polynucleotide sequences encoding	

CC	the antigen and the cytokine are under transcriptional control of same or		
CC	different promoter polynucleotide sequences. The expression vector, as a		
CC	DNA vaccine is useful for treating a condition in an organism. The		
CC	present sequence is human herpesvirus 2 complete DNA genome related		
CC	to the invention.		
xx			
50	Sequence 154746 BP; 23003 A; 54218 C; 54701 G; 22824 T; 0 other;		
Alignment Scores:			
Pred. No.:	621	Length:	154746
Score:	124.50	Matches:	84
Percent Similarity:	33.94%	Conservative:	28
Best Local Similarity:	25.45%	Mismatches:	116
Query Match:	8.62%	Indels:	102
DB:	24	Gaps:	13
US-09-245-198A-4 (1-284) x AAD25519 (1-154746)			
OY	10	AlaArgArgLeuProLeuProArgSerLeuGlySerArgAspGlyValAlaValArgIn	29
Db	32182	GGCGGTGTGGGCGCCACACAGCTGTCCCGCGGCCAGAGCGGTGTGGCTT	32123
OY	30	AlaGln-----ProProlAProMetAlAlaArgArgSerGlnArg	43
Db	32122	TCCCGCGCTGAGGAGACCCTTGGGCCACCGCGCCCGCTGTGCAACCTTGAACATGCG	32063
OY	44	ArgArgGly-----ArgArgGlyGluProGlyTThrAlaLeuLeuValProLeuAlaLeu	61
Db	32062	GGCCGAGGTGGCTTGTGTGCGCGCATCCAGATCGCGTGTGCGGCGCTTGGCCACTTC	32003
OY	62	GlyLeuGlyLeuAla-----LeuAlaCysLeuGlyLeuLeuLeu-----	74
Db	32002	TATCTGGGCGCGGTCTCCAGCGCGCCGCGCATCTTCCAGCGCTGCGGTAGCACTTGG	31943
OY	75	-----AlaValValSerLeuGlySerArgAlaSerLeuSerAlaGlnLeuProla	91
Db	31942	CCCCCACCAGACGACGATGGCTAGGGAGGTGACCGCGCTT-----CCCCGTCTCTGGA	31886
OY	92	GlnGlnGluLeuValAlaGlnGlnAspGlnAspProSerGluLeuAsnProGlnThrGlu	111
Db	31885	GAACCTGATGGCGGCCCTTCGAGGGGACCGC-----	31856
OY	112	GluSerGlnAspProlaIProPheLeuAsnArgLeuValArgProArgArgSerIaPro	131
Db	31855	-----CCCGACCGCTTCTTTCACACCGCGTACGCGCTTGGCGGTCTGGCGCA	31808
OY	132	LysGlyArgLysThrArgAlaArgAlaAlaAlaAla-----	144
Db	31807	TCTGGGCGGCCAGGCGCGGTGGGGCGCGCGCGCGCTGTCCCTTCAGGACGACAT	31748
OY	145	-----HisTryGlu	147
Db	31747	CCCCCGCGCTGCGCACCTCCAGCCACCACTACCGGTTGACTACTACAGCACACACCG	31688
OY	148	ValHisProArgProGlyGlnAspGlyAlaGlnAlaGly-ValAspGlyThrValSerG	167
Db	31687	AGACAGCGCTGGCGCTCCACCAACGCTCAATCGAGTGGATGACGGGAGCGTCAATGG	31628
OY	167	YTrPGlGlnAlaArgIleasn-----SerSerSerPro-----	178
Db	31627	ACGGCAGAGATMAATGTGCTTCATGGAGGCGACCCCTTCACCTGCGCCCAACAGGTT	31568
OY	179	-----LeuArgTryAsnArgI	184
Db	31567	CTGGCAGCACTACCTTCCGGGAGAGACGTACGCGTACCTCTCGCTTGGTTCAATCGCG	31508
OY	184	nileGlyLupheIleValIThrArgAlaGlyLeuTyTyrIleuTyTyrCysGlnValHisPh	204
Db	31507	CCCTCTGGGCGCTGTGTCTTCCGGGAGGCTTT-----GCCTT	31469
OY	204	eAspGluGlyLysAlaValTyrIleuLysLeuAspLeuValAspGlyValAlaLeuAla	224
Db	31468	TACCATTTACACCGCGCGTACGCTTCAGCGTCCAGACCCCGTCCGCGCTGCGGTGACT	31409

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 07:42:57 ; Search time 15.0648 seconds
(without alignments)
781.906 Million cell updates/sec

Title: US-09-245-198a-4
Perfect score: 1444
Sequence: 1 MSLLDFEISARRLPLPSRLG.....FWAHLKAPFLTYGFLFOVH 284

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Swissprot_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1268	87.8	249	1	TN12_HUMAN
2	1020	70.6	225	1	TN12_MOUSE
3	109.5	7.6	272	1	TN15_CHICK
4	108.5	7.5	260	1	TN15_CANFA
5	106.5	7.4	254	1	TN15_HUMAN
6	97	6.7	441	1	CG22_HUMAN
7	95.5	6.6	201	1	TN15_MOUSE
8	94	6.5	280	1	TN15_MOUSE
9	93.5	6.5	690	1	RHO_MICL
10	93	6.4	310	1	Y497_MYCTU
11	92	6.4	240	1	TN14_HUMAN
12	92	6.4	280	1	TN15_MOUSE
13	92	6.4	902	1	NRC4_HUMAN
14	90.5	6.3	760	1	MLH1_MOUSE
15	90.5	6.3	814	1	CADP_HUMAN
16	90	6.2	707	1	JIP1_MOUSE
17	89	6.2	280	1	MDCB_KLEPN
18	89	6.2	316	1	TN11_MOUSE
19	88.5	6.1	278	1	TN15_MOUSE
20	88.5	6.1	281	1	TN15_MOUSE
21	87	6.0	197	1	TN15_MOUSE
22	87	6.0	204	1	TN15_MOUSE
23	87	6.0	291	1	TN15_MOUSE
24	86.5	6.0	250	1	TN15_MOUSE
25	86	6.0	139	1	YOPB_BACSU
26	86	6.0	205	1	TN15_MOUSE
27	86	6.0	241	1	TN15_MOUSE
28	86	6.0	933	1	VGIB_HSV1
29	85.5	5.9	928	1	VGIB_HSV1
30	85.5	5.9	932	1	VGIB_HSV1
31	85	5.9	372	1	LKXB_MOUSE
32	85	5.9	379	1	LKXB_MOUSE
33	84.5	5.9	401	1	AROC_MYCTU

34	84	5.8	284	1	HXL1_HUMAN
35	84	5.8	310	1	TN15_MOUSE
36	84	5.8	575	1	MIS_PIG
37	84	5.8	825	1	ICP0_HSV2
38	83.5	5.8	416	1	RAGE_BOVIN
39	83.5	5.8	505	1	TUB_MOUSE
40	83.5	5.8	545	1	RTN2_HUMAN
41	83.5	5.8	885	1	VGIB_HSV2
42	83	5.7	228	1	BIOD_PSEAE
43	83	5.7	653	1	APPL_MOUSE
44	83	5.7	936	1	FHL1_YEAST
45	83	5.7	1217	1	AP4_MOUSE

ALIGNMENTS

RESULT 1	TN12_HUMAN	STANDARD:	PRT: 249 AA.
ID	043508: ORMU27;		
AC	15-JUN-2002 (Rel. 41, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Tumor necrosis factor ligand superfamily member 12 (TNF-related weak inducer of apoptosis) (TWEAK) (APO3 ligand).		
GN	TNFSF12 OR APO3L OR DR3LG.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A. AND N-TERMINUS OF SOLUBLE FORM.		
RC	TISSUE-Tonsil, and fetal liver.		
RX	MEDLINE=98070415; PubMed=9405449;		
RA	Chicopeite Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H., Hession C., Garcia I., Browning J.L.;		
RT	"TWEAK, a new secreted ligand in the tumor necrosis factor family that weekly induces apoptosis."		
RL	J. Biol. Chem. 272:32401-32410(1997).		
RM	[2]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-Fetal kidney;		
RX	MEDLINE=98228355; PubMed=9560343;		
RA	Martens S.A., Sheridan J.P., Pittl R.M., Brush J., Goddard A., Ashkenazi A.;		
RT	"Identification of a ligand for the death-domain-containing receptor APO3."		
RL	Curr. Biol. 8:525-528(1998).		
RM	[3]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE-Tonsil;		
RX	Submitted (Dec-2001) to the EMBL/GenBank/DBJ databases.		
RA	FUNCTION.		
RT	PubMed=10085077;		
RL	Lynch C.N., Wang Y.C., Lund J.K., Chen Y.-W., Leal J.A., Wiley S.R.;		
RM	"TWEAK induces angiogenesis and proliferation of endothelial cells."		
CC	J. Biol. Chem. 274:8455-8459(1999).		
CC	- FUNCTION: Binds to FN14 and possibly also to TNFRSF12/APO3. Weak inducer of apoptosis in some cell types. Mediates NF-kappaB activation. May promote angiogenesis and the proliferation of endothelial cells.		
CC	- SUBUNIT: HOMOTRIMER (POTENTIAL).		
CC	- TISSUE SPECIFICITY: Type II membrane protein and secreted. skeletal muscle, brain, colon, small intestine, lung, ovary, prostate, spleen, lymph node, appendix and peritoneal blood lymphocytes. Low expression in kidney, testis, liver, placenta, thymus and bone marrow. Also detected in fetal kidney, liver, lung and brain.		
CC	- PTM: The soluble form derives from the membrane form		

DS
SH1563

CC by proteolytic processing.

CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC -1- CAUTION: Ref.3 sequence differs from that shown due to a

CC frameshift in position 125.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AF030099; AAC51923.1; -

CC EMBL: AF055872; AAC39724.1; -

CC EMBL: BC019047; AAH19047.1; ALT_FRAME.

CC Genew: HGNC:11927; TNFSF12.

CC MIM: 602695; -

CC InterPro: IPR000478; TNF_family.

CC Pfam: PF00229; TNF; 1.

CC SMART: SM00207; TNF; 1.

CC PROSITE: PS00251; TNF_1; FALSE_NEG.

CC PROSITE: PS50049; TNF_2; 1.

CC CycloLine: Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.

CC CHAIN 1 249

CC CHAIN 94 249

CC FT DOMAIN 1 21

CC FT TRANSMEM 22 42

CC FT SITE 43 249

CC FT CARBOHYD 139 139

CC SEQUENCE 249 AA: 27216 MW: E660843361C28B8A CRC64:

Query Match 87.8%; Score 1268; DB 1: Length 249;

Best local Similarity 100.0%; Pred. No. 2e-93;

Matches 249; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

36 MARRSORRRGRGPGTALVPLALGLALACGLLALAVLSGRASISAOEPNOEL 95

1 MARRSORRRGRGPGTALVPLALGLALACGLLALAVLSGRASISAOEPNOEL 60

96 VAEEDDPSELNQTPEESODPAPFLNRLVPRRSAPKGRKTRARRAIAAYEHPRGOD 155

61 VAEEDDPSELNQTPEESODPAPFLNRLVPRRSAPKGRKTRARRAIAAYEHPRGOD 120

156 GAQAGVGVSGWEERINSSPLRNROIGETVTRAGLYLYCOVHEDEGKAVYIKLD 215

121 GAQAGVGVSGWEERINSSPLRNROIGETVTRAGLYLYCOVHEDEGKAVYIKLD 180

216 LAVDGVALARCEEFSAATASSLGPQLRCQVSGLLALRPGSSLRTRTPWALKKAPFL 275

181 LAVDGVALARCEEFSAATASSLGPQLRCQVSGLLALRPGSSLRTRTPWALKKAPFL 240

276 TYFGFLPOVH 284

241 TYFGFLPOVH 249

RESULT 2

TN12_MOUSE STANDARD: PRT: 225 AA.

AC 054907: 09CTP2:

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 12 (TNF-related weak

GN inducer of apoptosis) (TWEAK) (Fragment).

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI_TaxID=10090;

ON [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Peritoneal macrophage;

RX MEDLINE=98070415; PubMed=9405449;

RA Chicheportiche Y., Bourdon P.R., Xu H., Hsu Y.-M., Scott H.,

RA Hession C., Garcia I., Browning J.L.;

RT "TWEAK, a new secreted ligand in the tumor necrosis factor family that

RT weakly induces apoptosis.";

RL J. Biol. Chem. 272:32401-32410(1997).

RN [2]

RP SEQUENCE OF 83-225 FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=Retina;

RX MEDLINE=21085660; PubMed=11217651;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,

RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikiado I., Pesole G., Quackenbush J.,

RA Schriml L.M., Staudill F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamuya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,

RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

CC -1- FUNCTION: Binds to Fcγ14 and possibly also to TNFRSF12/Apo3. Weak

CC inducer of apoptosis in some cell types. Promotes angiogenesis and

CC the proliferation of endothelial cells. Mediates NF-κappa

CC activation (By similarity).

CC -1- SUBUNIT: Homotrimer (Potential).

CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By

CC similarity).

CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED.

CC -1- PTM: The soluble form is produced from the membrane form by

CC proteolytic processing (By similarity).

CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AF030100; AAC53517.1; -

CC EMBL: AK020909; BAB32249.1; -

CC MGD: MGI:1196259; Tnfsf12.

CC InterPro: IPR000478; TNF_family.

CC Pfam: PF00229; TNF; 1.

CC SMART: SM00207; TNF; 1.

CC PROSITE: PS00251; TNF_1; FALSE_NEG.

CC PROSITE: PS50049; TNF_2; 1.

CC CycloLine: Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.

CC NON_TER 1 225

CC CHAIN <1 225

CC FT CHAIN 70 225

CC FT TRANSMEM <1 21

CC FT SITE 69 70

CC FT DOMAIN 22 225

CC FT SITE 69 70

CC TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

CC MEMBER 12, MEMBRANE FORM.

CC TUMOR NECROSIS FACTOR LIGAND SUPERFAMILY

CC MEMBER 12, SECRETED FORM (BY SIMILARITY).

CC SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)

CC (POTENTIAL).

CC EXTRACELLULAR (POTENTIAL).

CC CLEAVAGE (BY SIMILARITY).

EMBO J. 13:616-624(1994).

-1- FUNCTION: ESSENTIAL FOR THE CONTROL OF THE CELL CYCLE AT THE G2/M (MITOSIS) TRANSITION. G2/M CYCLINS ACCUMULATE STEADILY DURING G2 AND ARE ABRUPTLY DESTROYED AT MITOSIS.

-1- SUBUNIT: INTERACTS WITH THE CDC2 AND CK2 PROTEIN KINASES TO FORM A SERINE/THREONINE KINASE HOLOENZYME COMPLEX. THE CYCLIN SUBUNIT IMPARTS SUBSTRATE SPECIFICITY TO THE COMPLEX.

-1- DEVELOPMENTAL STAGE: ACCUMULATES STEADILY DURING G2 AND IS ABRUPTLY DESTROYED AT MITOSIS.

-1- SIMILARITY: BELONGS TO THE CYCLIN FAMILY. CYCLIN AB SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; X76123; CAA53729.1; -

CC PIR; S41710; S41710.

CC HSSP; P30274; IYIN.

DR InterPro: IPR004366; Cyclin.

DR InterPro: IPR004367; Cyclin.Cterm.

DR Pfam; PF00134; cyclin; 1.

DR SMART; SM02984; cyclin_C; 1.

DR SMART; SM00385; CYCLIN; 2.

DR PROSITE; PS00292; CYCLINS; 1.

KW Cyclin; Cell cycle; Cell division; Mitosis.

SEQUENCE 441 AA; 49205 MW; E6E4C037C98880A7 CRC64;

Query Match 6.7%; Score 97; DB 1; Length 441;

Best Local Similarity 24.0%; Pred. No. 1.6;

Matches 70; Conservative 44; Mismatches 104; Indels 74; Gaps 15;

QY 25 GAVBOAPPPAPMAARRSOR-----RGRGEPGTLVPLALGLCLALACGL 73

DB 19 GAIRKQK-----MAYEKKRRALGIDGNVTVRGVKGKALPQVSRITRGF-----CAQLI 69

QY 74 -----LAVVSLGSRASLS-----AQEPQOEELVAEDDQPSLENPQTESQ 114

DB 70 ANAAAAAENNNKNSLAVNAKAGDGLPIKRAVARVPQKKYKSPQEIETISPTERRK 129

QY 115 DPAFLNLVLRPRS-----APKGRTRARALAAHYEVRPGQ-----DGAQAGVDCV 165

DB 130 --APVLEKEITGEKSLKKAAPLTSTLTARSKASV-VRTKPEQIYDIDADAVNNDLAV 186

QY 166 SGWEE-----ARINSSPLRY---NRQIEFIVTRAGLYLYCOVHFDP---EGKAVYL 212

DB 187 VEYEDMKFKFSAENDSRPHDYMSQPEINEMK--RAILLIDMLVQVHYKFELESFETLYL 244

QY 213 KLIDLVDGVALRC-----LEEPSATAASSLGPQLRLCOVSGLLAIRPGS 257

DB 245 TIN-IVDRYLTASKTTSRRELQLLGWSMLIAKYEIEWAPEVNDLVICISDGS 295

RESULT 7

TNFB_MACEU STANDARD: PRT; 201 AA.

ID TNFB_MACEU

AC Q9XT48;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Lymphotoxin-alpha precursor (LTA-alpha) (TNF-beta) (Tumor necrosis factor ligand superfamily member 1).

GN LTA OR TNFSF1 OR TNFB.

OS Macropus eugenii (Tamar wallaby).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Metatheria; Diprotodontia; Macropodidae; Macropus.

OC NCBI_TaxID=9315;

OX [1]

RN

RP SEQUENCE FROM N.A.

RX MEDLINE=20284948; PubMed=10826697;

RA Harrison G.A., Deane E.M.;

RT "cDNA cloning of lymphotoxin alpha (LT-alpha) from a marsupial, Macropus eugenii."

RT DNA Seq. 10:399-403(2000).

RL

CC -1- FUNCTION: Cytokine that in its homotrimeric form binds to TNFRSF1A/TNFR1, TNFRSF1B/TNFR and TNFRSF14/HVEM. In its heterotrimeric form with LTB binds to TNFRSF3/LTBR. Lymphotoxin is produced by lymphocytes and cytotoxic for a wide range of tumor cells in vitro and in vivo.

CC -1- SUBUNIT: Homotrimer, and heterotrimer of either two LTB and one LTA subunits or (less prevalent) two LTA and one LTB subunits (By similarity).

CC -1- SUBCELLULAR LOCATION: Secreted (homotrimer) and type II membrane protein (heterotrimers) (By similarity).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL; AF119336; AA041773.1; -

DR HSSP; P01374; TNF.

DR InterPro: IPR003636; TNF_abc.

DR InterPro: IPR000478; TNF_family.

DR Pfam; PF00228; TNF; 1.

DR PRINTS; PR01234; TNECROSISFCT.

DR PRODOM; PD002012; TNF_abc; 1.

DR SMART; SM00207; TNF; 1.

DR PROSITE; PS00251; TNF; 1.

DR PROSITE; PS50049; TNF_2; 1.

KW Cytokine; Glycoprotein; Cytotoxin; Signal.

FT SIGNAL 1 27

FT CHAIN 28 201

FT CARBOHYD 93 93

FT N-LINKED GLYCANC. (POTENTIAL).

SEQUENCE 201 AA; 21536 MW; 8C4C371C85091627 CRC64;

Query Match 6.6%; Score 95.5; DB 1; Length 201;

Best Local Similarity 23.2%; Pred. No. 0.86;

Matches 44; Conservative 27; Mismatches 84; Indels 35; Gaps 9;

QY 107 NPQTESODPAPFLRLVLRPRSAPKGRTRARAI--AAHYEVRPGQDGAQGVDCGT 164

DB 30 NPDNSHSSPAD-----PQRAHLQSKSLKREPLKRAHL-----VGDPSVQDSIH-- 75

QY 165 VSGWEARINSSP-LRYNRQI--GEFIVTRAGLYLYCOVHFDEGKA-----VYLKL 214

DB 76 ---W---RANIDHAFRLRIGFSLSNSLLVPTSGLFVYSQVVSAGASCSSEITPLLYLSH 129

QY 215 DLVDG---VIALRCLEEPSATAASSLGPQLRLCOVSGLLAIRPGSSLRIRLTPAHLKA 271

DB 130 EYLLFSSRYQVHVPFLISAKSVCSGTCQPMWRSYQGAVFLLTQGDRLSTYDGVSHLQ 189

QY 272 APFLTYFGLF 281

DB 190 SPSSVVFSGAF 199

RESULT 8

TNFB_MACMU STANDARD: PRT; 280 AA.

ID TNFB_MACMU

AC Q9BDM5;

DT 15-JUN-2002 (Rel. 41, Created)

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Tumor necrosis factor ligand superfamily member 6 (Fas antigen ligand) (CD95L protein).

GN TNFSF6 OR FASL OR CD95L.

OS Macaca mulatta (Rhesus macaque),

OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey), and

OS Macaca nemestrina (Pig-tailed macaque).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 OC NCBI_TaxID=9544, 9541, 9545;
 RN [1]
 RP SPECIES=M.mullata; TISSUE=Lymphocytes;
 RC MEDLINE=21383618; PubMed=1491535;
 RX Villingger F., Bostlik P., Mayne A.E., King C.L., Genain C.P.,
 RA Weiss W.R., Ansari A.A.;
 RT "Cloning, sequencing, and homology analysis of nonhuman primate
 Fas/FasLigand and co-stimulatory molecules.";
 RL Immunogenetics 53:315-328(2001).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC SPECIES=M.mullata, M. fascicularis, and M. nemestrina;
 RA Kiril'y., Inoue T., Yoshino K.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that
 transduces the apoptotic signal into cells. May be involved in
 cytotoxic T cell mediated apoptosis and in T cell development.
 CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of
 peripheral tolerance. In the antigen-stimulated suicide of mature
 T cells, or both. Binding to the decoy receptor TNFRSF6B/DCR3
 modulates its effects (by similarity).
 CC -1- SUBUNIT: Homotrimer (Potential).
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (by
 similarity).
 CC -1- PTM: The soluble form derives from the membrane form by
 proteolytic processing (by similarity).
 CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: AF344856; AAK37539.1; -;
 DR EMBL: AB035138; BAA90294.1; -;
 DR EMBL: AB035139; BAA90295.1; -;
 DR EMBL: AB035140; BAA90296.1; -;
 DR HSSP: P01375; 4TSV.
 DR InterPro: IPR003636; TNF-abc.
 DR InterPro: IPR000478; TNF-family.
 DR Pfam: PFO0229; TNF_1.
 DR PRINTS: PR01234; TNECROSISFCT.
 DR ProDom: PD002012; TNF-abc; 1.
 DR SMART: SM00207; TNF; 1.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 KM Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
 FT CHAIN 1 280
 FT
 FT CHAIN 129 280
 FT
 FT DOMAIN 1 80
 FT TRANSMEM 81 101
 FT
 FT DOMAIN 102 280
 FT
 FT DOMAIN 4 69
 FT SITE 128 129
 FT DISULFID 201 232
 FT CARBOHYD 183 183
 FT CARBOHYD 249 249
 FT CARBOHYD 259 259
 FT CONFLICT 60 60
 SO SEQUENCE 280 AA; 31367 MW; F0B284BD61A132EB4 CR64;

Query Match

6.5%; Score 94; DB 1; Length 280;

Best Local Similarity 20.8%; Pred. No. 1.6;
 Matches 62; Conservative 45; Mismatches 117; Indels 74; Gaps 14;
 QY 13 LPLPRISGSRDGAIVQAQAPAPMAARRSQRGRGEPPTALLVPLA----- 61
 DB 30 LPPPTSVPRRRGRRPPPPPPPP-----PPPSPLPPLPLPKKGNHST 78
 QY 62 GLGLALACGLLLAVVSLGSRASLSAEPAEELVAEEDDPSELNPQTEESDPAFLN 121
 DB 79 GLCLLVWFVFLVALVGLG--LGMFQLFHQKEL-----ALNESTSQKRTASLSEK 128
 QY 122 RLVPRRRSAPKGRKTRARRAIAAHYEHPRPGDGAQAGVDVGSQWEA-RINSSPLR 180
 DB 129 QIGHP---SPPEKKDQDK--VAHLTGKPMRSRSMPL-----WEDTYGIVLLSGVK 174
 QY 181 YNIOIGEFIVTRAGLVLYQCVHFDGKA-----VYLKID-----LVDGVLAAR 225
 DB 175 YKK--GGLVNETGLYFVSKYTF--RQOSTNLPLSHKRYMRNSKYPQDLVMEGKMST 231
 QY 226 CLEEFSTAAASLGLPDLRLCQVSGLLALRPSSLRITLPMHLKAPFLTYGRLPOV 283
 DB 232 CTTGQWMAHSSYLGAIVENLTSADHLV-----VNVSELVAFESQ--TFGLYKRL 280
 RESULT 9
 RHO_MICLU STANDARD; PRT; 690 AA.
 ID RHO_MICLU
 AC P52154;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Transcription termination factor rho.
 GN RHO.
 OC Micrococcus luteus (Micrococcus lysodeikticus).
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Micrococcales; Micrococcaceae; Micrococcus.
 OX NCBI_TaxID=1270;
 RN [1]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-5 AND 289-297.
 RC STRAIN-EM;
 RX MEDLINE=96132802; PubMed=8557681;
 RA Nowatzke W.L., Richardson J.P.;
 RT "Characterization of an unusual rho factor from the high G + C gram-
 RT positive bacterium Micrococcus luteus.";
 RL J. Biol. Chem. 271:742-747(1996).
 RN [2]
 RP SEQUENCE OF 205-690 FROM N.A.
 RC STRAIN-EM;
 RX MEDLINE=94327472; PubMed=8051015;
 RA Opperman T., Richardson J.P.;
 RT "Phylogenetic analysis of sequences from diverse bacteria with
 RT homology to the Escherichia coli rho gene.";
 RL J. Bacteriol. 176:5033-5043(1994).
 RN [3]
 RP REVISION TO 500.
 RC STRAIN-EM;
 RA Nowatzke W.L.;
 RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: FACILITATES TRANSCRIPTION TERMINATION BY A MECHANISM
 THAT INVOLVES RHO BINDING TO THE NASCENT RNA, ACTIVATION OF RHO'S
 RNA-DEPENDENT ATPASE ACTIVITY, AND RELEASE OF THE RNA FROM THE
 CC DNA TEMPLATE. RNA-DEPENDENT NPASE WHICH UTILIZES ALL FOUR
 CC RIBONUCLEOSIDE TRIPHOSPHATES AS WELL AS DAPF AS SUBSTRATES, BUT
 CC HAS A SIGNIFICANT LOWER ACTIVITY WITH CTP.
 CC -1- SUBUNIT: HOMOHXAMER (BY SIMILARITY).
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

RT "LIGHT, a new member of the TNF superfamily, and lymphotoxin alpha are
RT ligands for herpesvirus entry mediator."
RT Immunity 8:21-30(1998).
[2]
RN SEQUENCE FROM N.A. (ISOFORM 1), AND CHARACTERIZATION.
RC TISSUE-LIVER;
RX MEDLINE-98438532; PubMed-9765287;
RA Harrop J.A., McDonnell P.C., Brigham-Burke M., Lyn S.D., Minton J.,
RA Tan K.B., Dede K., Spemanato J., Silverman C., Hensley P.,
RA DiPonzio R., Emery J.G., Deen K., Eichman C., Chabot-Fletcher M.,
RA Truneh A., Young P.R.;
RT "Herpesvirus entry mediator ligand (HVEML), a novel ligand for
RT HVEML/RT, stimulates proliferation of T cells and inhibits HT29 cell
RT growth."
RT J. Biol. Chem. 273:27548-27556(1998).
[3]
RN SEQUENCE FROM N.A. (ISOFORM 2), AND PROCESSING.
RX MEDLINE-21528948; PubMed-11673523;
RA Granger S.W., Butrovich K.D., Houshamand P., Edwards W.R., Ware C.F.;
RT "Genomic characterization of LIGHT reveals linkage to an immune
RT response locus on chromosome 19p13.3 and distinct isoforms generated
RT by alternate splicing or proteolysis."
RL J. Immunol. 167:5122-5128(2001).
[4]
RN SEQUENCE FROM N.A.
RC TISSUE-Brain;
RA Strausberg R.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: Cytokine that binds to TNFRSF3/LTBR. Binding to the
CC decay receptor TNFRSF6B modulates its effects. Activates NFkB,
CC stimulates the proliferation of T cells, and inhibits growth of
CC the adenocarcinoma HT-29. Acts as a receptor for Herpes simplex
CC virus.
CC -1- SUBUNIT: Homotrimer.
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted
CC (isoform 1); Cytoplasmic (isoform 2).
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2/LIGHT
CC delta-TM; are produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: PREDOMINANTLY EXPRESSED IN THE SPLEEN BUT ALSO
CC FOUND IN THE BRAIN. WEAKLY EXPRESSED IN PERIPHERAL LYMPHOID.
CC TISSUES AND IN HEART, PLACENTA, LIVER, LUNG, APPENDIX, AND KIDNEY,
CC AND NO EXPRESSION SEEN IN FETAL TISSUES, ENDOCRINE GLANDS, OR
CC NONHEMATOPOIETIC TUMOR LINES.
CC -1- INDUCTION: UPREGULATED AFTER T-CELL ACTIVATION.
CC -1- PTM: N-glycosylated.
CC -1- PTM: The soluble form of isoform 1 derives from the membrane
CC form by proteolytic processing.
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -1- CAUTION: Ref.4 sequence differs from that shown due to a
CC frameshift in position 178.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF036581; AAC39563.1; -;
DR EMBL: AF064090; AAC25169.1; -;
DR EMBL: AY028261; AAK26160.1; -;
DR EMBL: BC018058; AAH18058.1; ALT_FRAME.
DR HSSP: P01375; 4TSV.
DR Genew: HGNC:11930; TNFSF14.
DR MIM: 604520; -;
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR000478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR ProDom: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF; 1.
DR PROSITE: PS00251; TNF_1; FALSE_NEG.
DR PROSITE: PS50049; TNF_2; 1.

KW Cytokine; Cytotoxin; Transmembrane; Glycoprotein; Signal-anchor;
KW Alternative splicing.
FT CHAIN 1 240
FT
FT CHAIN 783 240
FT
FT DOMAIN 1 37
FT TRANSMEM 38 58
FT
FT DOMAIN 59 240
FT SITE 82 83
FT DISULFID 154 187
FT CARBOHYD 102 102
FT VARSPLIC 38 73
FT CONFLICT 120 120
FT CONFLICT 214 214
SQ SEQUENCE 240 AA; 26351 MW; 49DDBF7EL390839 CRC64;

Query Match 6.4%; Score 92; DB 1; Length 240;
Best Local Similarity 23.7%; Pred. No. 2;
Matches 44; Conservative 19; Mismatches 57; Indels 66; Gaps 6;

QY 29 QAOPEPMARSRORRGEPGTALVPLALGLALACGLLAAVYSGSRASLSAQ 88
DB 16 QTDIPFTRIGSRHRQSCSVARVGLLL-LMGAGLAAGWFLQLHMRFG----- 66
QY 89 EPAQEEIYAEEDDPSELNPQTEESQDPAPFLNRLVPRRSAPKGRKTRARRAAHVEY 148
DB 67 -----EVV-----TRLPQAGSWEQLQGRSS-----HEV 92
QY 149 HPRPGDGAQAGVDGYSGWEARINSSPLRYNRQI-----GEIVTRAGLY 196
DB 93 NPAHLTGANSLSLTG-----SGGPLWETGLGLAFLRGLSYHDGALVTRAGY 141
QY 197 YLYCOV 202
DB 142 YLYSKV 147

RESULT 12
TNF6_CERTO
ID TNF6_CERTO STANDARD; PRT; 280 AA.
AC Q9BNI;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Tumor necrosis factor ligand superfamily member 6 (FAS antigen ligand)
DE (CD95L protein).
GN TNFSF6 OR FASL OR CD95L.
OS Cercopithecus torquatus alyx (Red-crowned mangabey) (Sooty mangabey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecus.
OX NCBI_TaxID=9531;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Lymphocytes;
RX MEDLINE-21383618; PubMed-11491535;
RA Villinger F., Bostlik P., Mayne A.E., King C.L., Genain C.P.,
RT Weiss W.R., Ansari A.A.;
RT "Cloning, sequencing, and homology analysis of nonhuman primate
RT Fas/FasL ligand and co-stimulatory molecules."
RL Immunogenetics 53:315-328(2001).
CC -1- FUNCTION: Cytokine that binds to TNFRSF6/FAS, a receptor that
CC transduces the apoptotic signal into cells. May be involved in
CC cytotoxic T cell mediated apoptosis and in T cell development.
CC TNFRSF6/FAS-mediated apoptosis may have a role in the induction of
CC peripheral tolerance. In the antigen-stimulated suicide of mature
CC T cells, or both. Binding to the decay receptor TNFRSF6B/DCR3
CC modulates its effects (by similarity).
CC -1- SUBUNIT: Homotrimer (Probable).
CC -1- SUBCELLULAR LOCATION: Type II membrane protein and secreted (By
CC similarity).

CC -1- PPM: The soluble form derives from the membrane form by
CC proteolytic processing (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TUMOR NECROSIS FACTOR FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF34847; AAK37606.1; -
DR HSSP: P01375; 4TSV.
DR InterPro: IPR003636; TNF_abc.
DR InterPro: IPR00478; TNF_family.
DR Pfam: PF00229; TNF_1.
DR PRINTS: PR01234; TNECROSISFCT.
DR ProDom: PD002012; TNF_abc; 1.
DR SMART: SM00207; TNF_1.
DR PROSITE: PS00251; TNF_1; 1.
DR PROSITE: PS50049; TNF_2; 1.
KW Cytokine; Apoptosis; Transmembrane; Glycoprotein; Signal-anchor.
FT CHAIN 1 280
FT DOMAIN 129 280
FT TRANSSEM 81 101
FT DOMAIN 1 80
FT TRANSSEM 81 101
FT DOMAIN 102 280
FT DOMAIN 4 69
FT SITE 128 129
FT DISULFID 201 232
FT CARBOHYD 183 183
FT CARBOHYD 249 249
FT CARBOHYD 259 259
SQ SEQUENCE 280 AA; 31407 MW; 729EA60067BD398 CRC64;

Query Match 6.4%; Score 92; DB 1; Length 280;
Best Local Similarity 20.8%; Pred. No. 2.4;
Matches 62; Conservative 44; Mismatches 118; Indels 74; Gaps 14;

QY 13 LPLRSISRGSGAVRQAQPPAPMAARSQRGRGEPGALLVPLAL----- 61
DB 30 LPCFTSVPRRRGQRRPPPPPPPPPP-----PPPPPLPPLPLPKKRGHNST 78
QY 62 GIGIALACIGLLAVNSIGSRASISAOEPAQELVAEEDODPSSELNPQTESODPAPFLN 121
DB 79 GLCLLVWFPMVLVALVGLG--LGMFQLFHLQKEL-----AELRESTSKHTASSLEK 128
QY 122 RLVPRRSAPKGRKTRARRAIAAHYEVHPRGQQAQGVDTVSGWEA-RINSSSPLR 180
DB 129 QIGHP--SPPEKKEQK--VAHLTKRNSRSMPL-----WEDTGIYVLISGVK 174
QY 181 YNRQIGETVRAGLYLYCOVHDEKA-----VYLKLD-----LLNGVYALR 225
DB 175 YKK--GGLVNTEGYFYYSKVF--RGQSCNTNPLSHKVMKNSKYPODLVMMBCKMMSY 231
QY 226 CLEFESATAASSLGPOLRICOVSGLLARPPSSSRIRTLPAHAHKAPFLTYGFLQV 283
DB 232 CTTCQMAHSHSYLGAVNLTSTHLY-----VNSLSLVNFEESG--TFGLYKL 280

DE transcription factor NFAT3 (NF-ATc4) (NF-AT3).
GN NFATC4 OR NFAT3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=T-cell; PubMed-7749981;
RX MEDLINE=95269130; PubMed-7749981;
RA Hoyer T., Sun Y.-L., Williamson K., Xu X.;
RT "Isolation of two new members of the NF-AT gene family and functional
RT characterization of the NF-AT proteins.";
RL Immunity 2:461-472(1995).
RN [2]
RP REVIEW.
RX MEDLINE=99189746; PubMed-10089876;
RA Crabtree G.R.;
RT "Genetic signals and specific outcomes: signaling through Ca2+,
RT calcineurin, and NF-AT";
RL Cell 96:611-614(1999).
CC -1- FUNCTION: PLAYS A ROLE IN THE INDUCIBLE EXPRESSION OF CYTOKINE
CC -1- GENES IN T CELLS, ESPECIALLY IN THE INDUCTION OF THE IL-2 AND IL-
CC 4 (BY SIMILARITY).
CC -1- SUBUNIT: MEMBER OF THE MULTICOMPONENT NFATC TRANSCRIPTION COMPLEX
CC THAT CONSISTS OF AT LEAST TWO COMPONENTS, A PRE-EXISTING
CC CYTOPLASMIC COMPONENT NFATC2 AND AN INDUCIBLE NUCLEAR COMPONENT
CC NFATC1. OTHER MEMBERS SUCH AS NFATC4, NFATC3 OR MEMBERS OF THE
CC ACTIVATING PROTEIN-1 FAMILY, MAP, GATA4 AND CBP/300 CAN ALSO BIND
CC TO THE COMPLEX. NFATC PROTEINS BIND TO DNA AS MONOMERS.
CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC FOR THE PHOSPHORYLATED FORM AND
CC NUCLEAR AFTER ACTIVATION THAT IS CONTROLLED BY CALCINEURIN-
CC MEDIATED DEPHOSPHORYLATION. RAPID NUCLEAR EXIT OF NFATC IS THOUGHT
CC TO BE ONE MECHANISM BY WHICH CELLS DISTINGUISH BETWEEN SUSTAINED
CC AND TRANSIENT CALCIUM SIGNALS. THE SUBCELLULAR LOCALIZATION OF
CC NFATC SPECIFY A KEY ROLE IN THE GENE TRANSCRIPTION.
CC -1- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN PLACENTA, LUNG, KIDNEY,
CC TESTIS AND OVARY. WEAKLY EXPRESSED IN SPLEEN AND THYMUS. NOT
CC EXPRESSED IN PERIPHERAL BLOOD LYMPHOCYTES.
CC -1- DOMAIN: REL SIMILARITY DOMAIN (RSD) ALLOWS DNA-BINDING AND
CC COOPERATIVE INTERACTIONS WITH AP1 FACTORS (BY SIMILARITY).
CC -1- PPM: PHOSPHORYLATED BY NFATC-KINASE; DEPHOSPHORYLATED BY
CC CALCINEURIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE REL/DORSAL FAMILY.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L41066; AAA79175.1; -
DR TRANSFAC: T02462; -
DR Genew: HGNC:7778; NFATC4.
DR MIM: 602699; -
DR InterPro: IPR002909; IPT_TIG.
DR InterPro: IPR000451; NF_Rel_dor_fam.
DR Pfam: PF01833; TIG; 1.
DR SMART: SM00429; IPT; 1.
DR PROSITE: PS01204; REL_1; FALSE_NEG.
DR PROSITE: PS50254; REL_2; 1.
KW Transcription regulation; Activator; Nuclear protein; DNA-binding;
KW Repeat; Phosphorylation.
FT DOMAIN 62 69
FT REPEAT 62 69
FT DOMAIN 114 119
FT DOMAIN 213 293
FT REPEAT 213 229
FT REPEAT 277 293
FT DOMAIN 297 304
FT DOMAIN 268 270
FT DOMAIN 430 437
POLY-PRO.
CALCINEURIN-BINDING.
2 APPROXIMATE SP REPEATS.
SP 1.
SP 2 (APPROXIMATE).
POLY-PRO.
NUCLEAR LOCALIZATION SIGNAL.
DNA-BINDING.

FT DOMAIN 672 674 NUCLEAR LOCALIZATION SIGNAL.
 SQ SEQUENCE 902 AA: 95472 MW: E59F15F7647A47C6 CRC64;
 Query Match 6.4%; Score 92; DB 1; Length 902;
 Best Local Similarity 25.0%; Pred. No. 9; Mismatches 73; Indels 56; Gaps 8;
 Matches 48; Conservative 15; Mismatches 73; Indels 56; Gaps 8;
 QY 14 PLPRSLGSRDGCAGVROAPAPMAARRS-----QRRGRGEPDTALLVPLALGLALAC 69
 DB 236 PPSGGRGPDSDMLLSAPGPTASPRPAPSGKRRYSSSGTPSSA-----SPALGR 286
 QY 70 LGILLAVSLGSRASLSAOPAEELVAEEDDPSSELNQTESQDPAPPLNRLVPRRS 129
 DB 287 RG-----SLGEGS-----EPPPPPL-PLARDPGSGPFDYVGAPPAES 325
 QY 130 AP-KGRKTRARRAIA-----AHYEVHPRPGDGAQAGVD-----GTVSG 167
 DB 326 IQQKTRRTSSQVAVLPREEEPASCKLPLGAESVAPPGGSRKKEVAGMDYLAVPSPLA 385
 QY 168 WEARINSSSPL 179
 DB 386 WSKARIGCHSPT 397
 RESULT 14
 MLH1_MOUSE STANDARD: PRT: 760 AA.
 ID MLH1_MOUSE 062454;
 AC 09JRK91: 062454;
 DT 16-OCT-2001 (rel. 40, Created)
 DT 16-OCT-2001 (rel. 40, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE DNA mismatch repair protein Mlh1 (Mult. protein homolog 1).
 GN MLH1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue-Testis;
 RA Kumaran M., Rao M.R.S.;
 RT "Cloning of the cDNA of the mult. homolog, MLH1 from mouse testis."
 RL Submitted (Mar-2000) to the EMBL/Genbank/DBJ databases.
 RP SEQUENCE OF 1-151 FROM N.A.
 RX MEDLINE=96270514; PubMed=8674118;
 RA Edelmann W., Cohen P.E., Kane M., Lau K., Morrow B., Bennett S.,
 RA Koldner R.D., Kuchelapati R., Chaganti R., Pollard J.W.,
 RT "Meiotic pachytene arrest in Mlh1-deficient mice."
 RL Cell 85:1125-1134(1996).
 CC -1- FUNCTION: PROBABLY INVOLVED IN THE REPAIR OF MISMATCHES IN DNA.
 CC -1- SUBUNIT: HETERODIMER OF MLH1 AND PMS2 OR MLH1 AND MLH3 (BY
 CC SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE DNA MISMATCH REPAIR MUTL/HEXB FAMILY.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC at the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AF250844; AAF64514.1; -
 CC EMBL: U60872; AAC52672.1; -
 CC EMBL: U59881; AAC52672.1; JOINED.
 CC EMBL: U59883; AAC52672.1; JOINED.
 CC EMBL: U59884; AAC52672.1; JOINED.
 CC HSSP: P23367; LBKN.
 DR MGD: MGI:101938; Mlh1.
 DR InterPro: IPR003594; ATPbind_ATPase.

DR InterPro: IPR002099; DNA_mis_repair.
 DR Pfam: PF01119; DNA_mis_repair; 1.
 DR Pfam: PF02518; HATPase_C; 1.
 DR TIGRPFAM: TIGR00585; mutL; 1.
 DR PROSITE: PS00058; DNA_MISMATCH_REPAIR_1; 1.
 KW DNA repair.
 SQ SEQUENCE 760 AA: 84679 MW: 173C809372A29186 CRC64;
 Query Match 6.3%; Score 90.5; DB 1; Length 760;
 Best Local Similarity 22.7%; Pred. No. 9, 8;
 Matches 70; Conservative 31; Mismatches 126; Indels 81; Gaps 14;
 QY 13 LPLPRSLGSRDGCAGVRO-----AOPAPMAARRSQRRGRGEPGALLVPLALGLG 64
 DB 436 LPAPAEAAESENLERESIMETSDAAQKAAPTSSPSSRRHR--EDSDVENENASGKE 493
 QY 65 LALAC-----LGILLAVSLGSRASLSAOPAEELVAEEDDPSSELNQ--TEESQDRA 117
 DB 494 MTAACYPRRRRIINLTSLQOEISERCHETURE--ILRNHSFVGCVNQWALAOHOTKL 551
 QY 118 PFLNRLVRRRSAPKGRKTRARRAIAHYEV-----HPRPGDGAQAGVDG 163
 DB 552 YLLN-----TTKLSEELFYQILLYDFRANGVLRSEPPAPFLDLAMALDS 596
 QY 164 TVSGWEARINSSSPLRYNRQIGFIVTRAGLYLYICQVHDEGKAVYLUKLDLYDG--- 220
 DB 597 PESGWTEDDQPKEGLAEV---IVEFLKKAEMLADYFSEIDE--EGNLIGLPLLDISYVP 652
 QY 221 -----VLALR-----CLEEFSATAAS--SLGPQ--LRICQVSGILLALRPGSS 258
 DB 653 PLEGPLIFLRLATEVNMDEKEKCFESLSKECAMFYSIKROYLIEBSTLSGQSDMPGST 712
 QY 259 LRIRTLPW 266
 DB 713 SK-----PW 716
 RESULT 15
 CADF_HUMAN STANDARD: PRT: 814 AA.
 ID CADF_HUMAN P55291;
 AC 01-OCT-1996 (rel. 34, Created)
 DT 01-OCT-1996 (rel. 34, Last sequence update)
 DT 16-OCT-2001 (rel. 40, Last annotation update)
 DE Muscle-cadherin precursor (M-cadherin) (Cadherin-15) (Cadherin-14).
 GN CDH15 OR CDH14 OR CDH3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC Tissue-Skeletal muscle;
 RA MEDLINE=97184182; PubMed=9030594;
 RX Shibata T., Shimoyama Y., Goton M., Hirohashi S.;
 RT "Identification of human cadherin-14, a novel neurally specific type
 RT II cadherin, by protein interaction cloning."
 RL J. Biol. Chem. 272:5236-5240(1997).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC Tissue-Muscle;
 RA Strausberg R.;
 RT Submitted (May-2001) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE
 CC SORTING OF HETEROGENEOUS CELL TYPES. M-CADHERIN IS PART OF THE
 CC MYOGENIC PROGRAM AND MAY PROVIDE A TRIGGER FOR TERMINAL MUSCLE
 CC DIFFERENTIATION.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: March 31, 2003, 01:56:48 ; Search time 1878.64 Seconds

(without alignments)
2448.323 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARLRLPRSLG.....FWAHLKAPFLTYGFLQVH 284

Scoring table:

BIOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Delop 6.0, Delext 7.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n model -DEV=xlh
-O=/qgr2.1/USPRO/spool/US09245198/runat_24032003.163554.25956/app.query.fasta.1.846
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -DOPEXT=0
-UNITS-bits -START=1 -END=-1 -MATRIX=BIOSUM62 -TRANS=human40.cdi -LIST=45
-DOCLIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=plc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09245198 @CGN.1.1.2463 @runat.24032003.163554.25956 -NCPU=6 -ICPU=3
-NO_XLPHY -NO_MMAP -LARGQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: em_estba:*
2: em_esthum:*
3: em_estlin:*
4: em_estlm:*
5: em_estov:*
6: em_estpl:*
7: em_estio:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vit:*
22: em_gss_fun:*
23: em_gss_man:*
24: em_gss_mus:*
25: em_gss_pio:*
26: em_gss_rtd:*
27: em_gss_rtd:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Length	DB ID	Description
1	1087.5	75.3	731	13	BI871711
2	1060.5	73.4	728	13	BI870393
3	979	67.8	777	13	BI819200
4	962	66.6	834	13	BI766766
5	934	64.7	828	13	BI596681
6	913	63.2	918	13	BE577781
7	908.5	63.0	948	14	EO707185
8	904	62.6	940	14	BO884231
9	825	57.1	561	10	AW763237
10	809	56.0	963	14	BO671259
11	765	53.0	1071	14	BM921213
12	753	52.1	951	14	BO674188
13	742.5	51.4	785	13	BI762908
14	738	51.1	587	12	BG686319
15	707	49.0	1033	11	AK020909
16	652	45.2	584	10	AW917574
17	648.5	44.9	698	13	BI906850
18	647	44.8	894	13	BI908274
19	645.5	44.7	474	13	BI965174
20	604	41.8	445	9	AA870722
21	598.5	41.4	471	9	AA221610
22	582.5	40.3	542	12	BF041509
23	582	40.3	650	12	BG404836
24	572.5	39.6	404	12	BF044430
25	562	38.9	360	10	BE654876
26	520.5	36.0	493	10	BE307031
27	506	35.0	538	12	BF821434
28	499	34.6	558	13	BM484863
29	493	34.1	298	14	BM688946
30	476	33.0	441	13	BI967060
31	466	32.3	315	12	BF466521
32	441	30.5	465	9	AT091441
33	440	30.5	581	13	BI738634
34	421	29.2	345	14	RS5379
35	420	29.1	278	10	AA417023
36	403	27.9	1187	14	BO053284
37	395.5	27.4	571	12	BF073881
38	391	27.1	639	10	BB642326
39	388.5	26.9	1926	14	BM906056
40	363	25.1	400	9	AI152313
41	362	25.1	342	9	AA637970
42	353	24.4	910	12	BG110063
43	352	24.4	456	13	BI966235
44	347	24.0	487	10	AM320117
45	320	22.2	533	10	BE628951

ALIGNMENTS

RESULT 1
BI871711
LOCUS 603395825P1 NIH_MGC_90 Homo sapiens CDNA clone IMAGE:5405478 5',
DEFINITION mRNA sequence.
ACCESSION BI871711 731 bp mRNA
VERSION BI871711.1 GI:16045386
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 731)
NIH-MGC http://mgc.nci.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE

JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph. D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12034 row: b column: 07
High quality sequence stop: 728.
Location/Qualifiers
1. 731
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5405478"
/clone_1lb="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 125 a 242 c 228 g 136 t
ORIGIN
Alignment Scores:
Pred. No.: 2,28e-93 Length: 731
Score: 1087.50 Matches: 219
Percent Similarity: 99.10% Conservative: 0
Best Local Similarity: 99.10% Mismatches: 1
Query Match: 75.31% Indels: 2
DB: 13 Gaps: 1
US-09-245-198a-4 (1-284) x B1871711 (1-731)
QY 65 LeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSer 84
|||||
Db 2 CTGGCGCTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 61
QY 85 LeuSerAla---GlnGluProAlaGlnGluGluLeuValAlaGluGluAspGlnAspPro 103
|||||
Db 62 CTGTCGCCCGCAG 121
QY 104 SerGluLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeu 123
|||||
Db 122 TCGGAAGTGAATCCCGCAGACAGAAAGCCAGAGATCTCGCTTCCGACGACGACTA 181
QY 124 ValArgProArgSerAlaProGlyGlyArgLysThrArgAlaArgArgAlaIleAla 143
|||||
Db 182 GTTCGGGCTCGCAGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 241
QY 144 AlaHisTyrGluValHisProArgProGlyGlnAspGlyValAlaGlyValAspGly 163
|||||
Db 242 GCCCATTTATAGTTCTCCACGACCTGGACAGAGAGAGAGAGAGAGAGAGAGAGAG 301
QY 164 ThrValSerGlyTyrGluGluAlaArgIleAsnSerSerSerProLeuArgTyrAsnArg 183
|||||
Db 302 ACGATGAGTGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 361
QY 184 GlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrGlnValHis 203
|||||
Db 362 CAGATCGGGAGTTTATAGTCACCCGGGCTGCTACTACTGTAAGTCAAGTGCAC 421
QY 204 PheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGlyValLeuAla 223
|||||
Db 422 TTTGATGAGGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 481
QY 224 LeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArg 243
|||||
Db 482 CTGCGCTGCTGGAGAAATTTCTAGCCACTGCGGCGAGTTCCCTCGGGGCCCGCAGCTCGC 541

QY 244 LeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArgThr 263
|||||
Db 542 CTGCGCAGAGTGTCTGGGCGCTGGGCGCTGGGCGGACAGCTGC-TCCCTGGGAGATCCGACG 600
QY 264 LeuProTrrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnVal 283
|||||
Db 601 CTCGCCCGGGGCCCATCTCAAGGCTGGCCCTTCTCTCACTTCTGCGACTTCCACGTT 660
QY 284 His 284
|||
Db 661 CAC 663
RESULT 2
B1870393 728 bp mRNA linear EST 11-OCT-2001
LOCUS 603395641F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:5405459 5',
DEFINITION mRNA sequence.
ACCESSION B1870393 GI:16044066
VERSION B1870393
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 728)
NIH-MGC http://mgc.ncl.nih.gov/.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL
COMMENT Contact: Robert Strausberg, Ph. D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12034 row: a column: 12
High quality sequence stop: 728.
Location/Qualifiers
1. 728
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5405459"
/clone_1lb="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 125 a 240 c 227 g 136 t
ORIGIN
Alignment Scores:
Pred. No.: 8.44e-91 Length: 728
Score: 1060.50 Matches: 218
Percent Similarity: 98.64% Conservative: 0
Best Local Similarity: 98.64% Mismatches: 2
Query Match: 73.44% Indels: 3
DB: 13 Gaps: 1
US-09-245-198a-4 (1-284) x B1870393 (1-728)
QY 65 LeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValSerLeuGlySerArgAlaSer 84
|||||
Db 2 CTGGCGCTGGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCG 61
QY 85 LeuSerAla---GlnGluProAlaGlnGluGluLeuValAlaGluGluAspGlnAspPro 103
|||||
Db 62 CTGTCGCCCGCAG 121

OY	104	SerGIuLeuAaPrOcIInThGInuSerGIaNaPrOcIaAProPhetLeuAaSrGIeU	123
Db	122	TCGGAGATGAATCCCAAGACAAAGAACCAAGATCTCGCCCTTTCTCGTAACCGACTA	181
OY	124	ValAaPrOcIaAaSerGIaAProLySGIaAqLySThrArGIaAArgaAlaIleAa	143
Db	182	GTTCGGCCTCCGAGAAAGTGCACCTTAAAGCCGGAACACAGGGCTCTGAAGACGATCCGA	241
OY	144	AlaHISrYrGIuValaHISProArGIProGIyGIaNaSPGIyAlaGInaIaGlyValaSPGIy	163
Db	242	GCCCATTTATGAAGTTCATCCACGACCTGGACAGCAGGAGCCGACGAGTGTGCACGGG	301
OY	164	ThrValSerGIyTPrGIuGInaIaArgIleAaSrSerSerProLeuAaArgTyrAaArg	183
Db	302	ACAGGAGATGTGGTGGAGAGAACCGAATCAACACACTCCAGCCCTTCGCTACAAACCGC	361
OY	184	GInIleGIyGIuPhetIleValThrArGIaGIaGlyLeuTyrTyrLeuTyrCysGInaIaHIS	203
Db	362	CAGATCGGGGAGATTATATAGTCAACCGGGCTGGGGCTCTACTACTGTAGGTGCAC	421
OY	204	PheaSPGIuGIyLysAlaValaTyrLeuLysLeuAaSPLeuLeuValaSPGIyValaLeuAa	223
Db	422	TTTGATGAGGGGAAAGCTGTCTACTGAAGCTGAGACTGTGCTGATGTGTCTGGCC	481
OY	224	LeuArGIyGIeGIuGIuPhetSerIaHrAlaAlaSerSerLeuGIyProGIleuAaArg	243
Db	482	CTGGCTCTCTCGAGGAATCTTCACGCACCTGGGACAGTCCCTCGGGCCAGCTCCGC	541
OY	244	LeuCysGIaIaSerGIyLeuLeuAlaLeuAaPrOGIySerSerLeuAaArgIleArGIhr	263
Db	542	CTCTGCCAAGTGTCTGGGGCTGTGGCCCTGGCGCCAGAGTCC-TCCCTGGGATCCGCACAC	600
OY	264	LeuPrOTPrAlaHISLeuLysAlaAlaPrOheLeuThrTyrPhetGIyLeuPhetGIaIa	283
Db	601	CTCCCCCTGGGCCAATCTCAAGGCTGC-CCCTTCCTCACCCTACTCTGAGACTCTCCAGGTT	659
OY	284	HIS 284	
Db	660	CAC 662	
RESULT 3			
LOCUS	B1819200	777 bp	mRNA linear EST 04-OCT-2001
DEFINITION	60303461Arl NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5175658 5',		
ACCESSION	B1819200		mRNA sequence.
VERSION	B1819200.1	GI:15930750	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homiini; Homiidae; Homo.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/ .		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgabbs-remail.nih.gov		
	Tissue Procurement: Life Technologies, Inc.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/MLN at:		
	http://image.llnl.gov		
	Plate: L14M11437 row: 1 column: 03		
	High quality sequence stop: 759.		
FEATURES	Location/Qualifiers		
source	1..777		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:5175658"		

Accession	Sequence	Gene	Species	Clone	Image
US-09-245-198A-4 (1-284)	x B1819200 (1-777)				
QY	97	AlaGluGluAspGluAspProSerGluLeuAsnProGlnThrGluGluSerGlnAspPro	116		
Db	14	GCAGAGGAGGACGACGAGCCCGTCGGAGACTGATGCCACAGAGAAAGCAGATCCT	73		
QY	117	AlaProPheLeuAsnArgLeuValArgProAlaArgSerAlaProGlyArgGlyThr	136		
Db	74	GGCGCTTCCTGTAACCGACTAGTTGGGCTCCCAAGAGCCACCTAAAGCCCGGAAACA	133		
QY	137	ArgAlaArgArgAlaIleAlaIleHisTyrGluValHisProArgProGlyGlnAspGly	156		
Db	134	CGGGCTCGAAGAGCCATCGCAGCCCATTTATGAATTCATCCAGCAGCTGGACAGACGGA	193		
QY	157	AlaGlnAlaGlyValAspGlyThrValSerGlyTyrPrlGluAlaIleArgIleAsnSerSer	176		
Db	194	GGCGAGGACGAGTGTGGAGCGGACAGTGAAGTGGCTGGGAGGAAGCCAGATCAACAGCTCC	253		
QY	177	SerProLeuArgTyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyr	196		
Db	254	ACCCCTCTGGCTACACCGCCAGATCGGGAGTTAATAGTACCCGGGCTGGCTTAC	313		
QY	197	TyrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTyrLeuLysLeuAspLeu	216		
Db	314	TACCTGTACTGTCAAGTGTGACTTGTATGTAGGGGAAAGGCTGTACTGTGAAGCTGACTTG	373		
QY	217	LeuValAspGlyValLeuAlaLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaIleSer	236		
Db	374	CTGGTGGAGATGATGTCTGTGGCCCTGGGCTGCCCGAGGAATTCACAGCCATCGGCGAGT	433		
QY	237	SerLeuGlyProGlnLeuArgLeuGlyCysGlnValSerGlyLeuLeuAlaLeuArgProGly	256		
Db	434	TCCCTCTGGGCCCCACACTCCGCTTGCACAGGTGTGTGGGCTTGTGGCCCTGGGCGCAGAG	493		
QY	257	SerSerLeuArgIleAlaArgThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThr	276		
Db	494	TCCCTCCCTGGGATCGGACCCCTCCCTGGGCGCACATCTCAAGCTGACCCCTCTCCACCC	553		
QY	277	TyrPheGlyLeuPheGlnValHis	284		
Db	554	TACTTCGGAGCTTCCAGGTTTAC	577		
RESULT 4					
LOCUS	B1766766	834 bp	mRNA	linear	EST 25-SEP-2001
DEFINITION	603056666r1 NIH_MGC_122	Homo sapiens	cdna	clone IMAGE:5206217	5',
ACCESSION	B1766766				
VERSION	B1766766.1	GI:15758344			
KEYWORDS	EST				

```

SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 834)
AUTHORS     NIH-MGC http://mgi.ncl.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-remail.nih.gov
            Tissue Procurement: Life Technologies, Inc.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM11517 row: c column: 18
            High quality sequence stop: 772.
            Location/Qualifiers
                1..834
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5206217"
                /clone_lib="NIH_MGC_122"
                /lab_host="DH10B"
                /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
                Site_1: NotI; Site_2: EcoRV (destroyed); RNA source:
                anonymous pool of 24 week female lung, 16 week female
                spleen, and 20-22 week male spleen. Library is oligo-dT
                primed and directionally cloned (EcoRV site is destroyed
                upon cloning). Average insert size 1.4 kb, insert size
                range 1-3 kb. Library is normalized and enriched for
                full-length clones and was constructed by C. Gruber
                (Invitrogen). Research Genetics tracking code 026. Note:
                this is a NIH-MGC Library."
BASE COUNT  154 a      287 c      223 g      170 t
ORIGIN
Alignment Scores:
Pred. No.:      2,46e-81      Length:      834
Score:          962.00        Matches:      192
Percent Similarity: 98.97%      Conservative: 0
Best Local Similarity: 98.97%      Mismatches: 1
Query Match:    66.62%      Indels:      2
DB:             13          Gaps:      0

US-09-245-198a-4 (1-284) x B1766766 (1-834)
OY      92  GlnGluGluLeuValAlaGluGluAspGluAspProSerGluLeuAsnProGlnThrGlu 111
         |||||||
Db      3  CAGAGAGAGAGCTGTGCGACAGAGAGACAGAGACCCGTCGAGACTGAATCCCGACAGAA 62
OY      112 GluSerGlnAspProAlaProPheLeuAsnArgLeuValArgProArgSerAlaPro 111
         |||||||
Db      63  GAAAGCCAGAGATCCCTGCGCTTCGAAACCGACTGAGTGGCTCCGACAGTGCACCT 122
OY      132 LysGluArgPylThrArgAlaArgAlaIleAlaAlaHisThrGluValHisProArg 151
         |||||||
Db      123  AAAGGCCGGAACACAGGCTCGAAGAGCGATGCGACCCATATGAAAGTTCCACGA 182
OY      152 Pro-GluGlnAspGlyAlaGlnAlaGlyValAspGlyThrValSerGlyTrpGluGlnAl 171
         |||||||
Db      183  CCGTGGACAGAGAGGAGCGAGCGAGGTGTGAGACGGGACAGTGAAGTGGCGAGAGAGC 242
OY      171  ArgGluLeuAsnSerSerSerProLeuArgTyrArgAsnArgGlnIleGlyGluPheIleValTh 191
         |||||||
Db      243  CAGAATCAACAGAGCTCCAGCCCTCTGCGCTACAAACCGCCAGATGGGGAGTTATAGCAC 302
OY      191  ArgAlaGlyLeuLeuTyrThrLeuTyrCysGlnValHisPheAspGluGlyValAlaValTy 211
         |||||||
Db      303  CCGGGCTGGGCTTACTACTGACTGTGACAGGTGACCTTTGATGAGGGAGGCGTGTCTA 362

```

```

OY      211  rLeuLysLeuAspLeuLeuValAspGlyValAlaGlnAlaLeuArgCysLeuGluGluPheSe 211
         |||||||
Db      363  CTTGAAGCTGAGACTGCTGTGATGATGCTGCGCCCTGCTGCTGCGAGAAATTTCT 422
OY      231  rAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuGlyGlnValSerGlyLeuLe 251
         |||||||
Db      423  AGCCACTGCGGCGAGTTCCTCGGGCCCGACGCTCCGCTCTGCGAGAGTGTGGGCTTT 482
OY      251  uAlaLeuArgProGlySerSerLeuArgIleArgThrLeuProTrpAlaHisLeuLysAl 271
         |||||||
Db      483  GGGCCCTGCGGCGCAGGTTC-TCCTGGGGATCGCACCCCTCCCTGGGGCCATCTCAAGGC 541
OY      271  aAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis 284
         |||||||
Db      542  TGGCCCTTCCTCCACTACTTCGAGCTTCCAGGTTTAC 581
RESULT 5
LOCUS      B1596681      828 bp      mRNA      linear      EST_07-SEP-2001
DEFINITION 603243254F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5285892 5'',
            mRNA sequence.
ACCESSION  B1596681
VERSION    B1596681.1 GI:15489620
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 828)
AUTHORS     NIH-MGC http://mgi.ncl.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgaabs-remail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroaki
            Toshiyuki and Piero Carninci (RIKEN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM11722 row: k column: 13
            High quality sequence stop: 776.
            Location/Qualifiers
                1..828
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:5285892"
                /clone_lib="NIH_MGC_96"
                /tissue_type="hypothalamus"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pBluescript (modified
                pluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag
                ); Oligo-dT primed using primer 5'-TTTCTTTTCTTTTCTTTT-3',
                size-selected for average insert size 2.3 kb and
                normalized to R0T 5. This is a primary library enriched
                for full-length clones and constructed using the
                Cap-trapper method (Carninci, in preparation). Library
                constructed by M. Brownstein (NHGRI/NHGRI, National
                Institutes of Health). Note: this is a NIH-MGC Library."
BASE COUNT  155 a      278 c      223 g      172 t
ORIGIN
Alignment Scores:
Pred. No.:      1.13e-78      Length:      828
Score:          934.00        Matches:      179
Percent Similarity: 100.00%      Conservative: 1
Best Local Similarity: 99.44%      Mismatches: 0
Query Match:    64.68%      Indels:      0
DB:             13          Gaps:      0

US-09-245-198a-4 (1-284) x B1596681 (1-828)

```


Oy		105	GlutEuaAnPrgInThrGluGlnSerClnsPrAlaProPhleAAsnArgEuaVal	124
Db		44	GAACTGAATCCCGACAGAAAGAAGCCAGACTCCGGCTTCTTAACCAGACTAGTT	103
Oy		125	ArgProAgaCgSerAlaProLysGlyArgrYrThraGlaIarArgAlaIlleaAla	144
Db		104	CGGGCTGCACAAGTCACCTAAAGGCCGGAAGAACAGGGCTCGAAGAGCATGCCAAC	163
Oy		145	HlsYrrGluValHlsProArGrProGlyClnAspgLyAlaGlnAlaGlyValAspgLyThr	164
Db		164	CATTATGAAGTTCATCCACGACCTGGACAGACGACGACGCGAGCGTGCGACGGGACA	223
Oy		165	ValSerGlyTPPgUgluAlaArGrIleAsnsSerSerProLeuArGrYrAsnArgGln	184
Db		224	TTCAGTGCCTGGAGAAAGCCAGAAATCAACAGCTCCAGCCCTCTGCGCTACACCGCAG	283
Oy		185	IleaglIupheIleValThrArGaIagLyLeuTyrrYrLeuTyrcysGlnValHisphe	204
Db		284	ATCGGGAGATTATATAGTACCCCGGGCTGGGCTCTACTACTGTACTGTACAGTGACATT	343
Oy		205	AspGluGlyLysAlaValTyrIleuLysLeuAspleuValAspgLyValleuAlaLeu	224
Db		344	GATGGGGGAAAGCGCTGTACTCTGAAGCTGACCTTGCGTAGGAGTGCTGTGCGCTG	403
Oy		225	ArgCysLeuGluGluPhSerAlaThrAlaIaAserSerLeuGlyProGlnleuArgrLeu	244
Db		404	CGCTGCCCTGGAGAAATTCACCCACCTCGGCCACTTCCCTCGGGCCCCAGCTCGCCTC	463
Oy		245	CysGlnValSerGlyLeuAlaLeuValArgProGlySerSerLeuArgIleArgrThrLeu	264
Db		464	TGCCAGGTCTGTGGGCTTGTGGCCCTCGCGGCGAGGGTCTCCCTCGGATCGCACCTTC	523
Oy		265	ProTrpAlaHlsIeuLysAlaAlaProPhleuThrYrIrrPhesGlyLeuPhcGlnValHis	284
Db		524	CCCTGGGCCCACTCAAGGCTGCCCCCTTCTCCACTCACTTCGAGCTCTCCAGGTTCAC	583
RESULT 6				
LOCUS	Bf577781			
DEFINITION	602092080P1 NCL_CGAP_CO24 Mus musculus cDNA IMAGE:4206595 5'	918 bp	mRNA	linear EST 12-DEC-2000
ACCESSION	Bf577781			
VERSION	Bf577781.1			
KEYWORDS	EST.			
SOURCE	house mouse.			
ORGANISM	Mus musculus			
REFERENCE	EuaAyVol; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.			
TITLE	1 (bases 1 to 918)			
JOURNAL	NIH-MSC http://mgc.nci.nih.gov/.			
COMMENT	National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cgabbs@email.nih.gov Tissue Procurement: Jeffrey E. Green, M.D. cDNA Library Preparation: Life Technologies, Inc. cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: LHAM9767 row: P column: 20 High quality sequence start: 17 High quality sequence stop: 724. Location/Qualifiers 1..918 /organism="Mus musculus" /strain="FVB/N" /db_xref="taxon:10090" /clone="IMAGE:4206595" /clone_lib="NCL_CGAP_CO24" /lab_host="DH10B (TI phage-resistant)".			
FEATURES				
SOURCE				

/note="Organ: colon. Vector: PCMV-SpORT6; Site:1: NotI; Site_2: SalI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.6 kb. Constructed by life technologies. Note: this is a NCI/CGAP Library."

BASE COUNT	153 a	292 c	268 g	204 t	1 others
ORIGIN					

Alignment Scores:	1,31e-76	Length:	918
Pred. No.:	913.00	Matches:	191
% Score:	91.74%	Conservative:	9
Percent Similarity:	91.74%	Mismatches:	18
Best Local Similarity:	63.23%	Indels:	3
Query Match:	12	Gaps:	0

US-09-245-198A-4 (1-284) x BF577781 (1-918)

QY	67	LeuAlaGysLeuGlyLeuLeuLeuAlaValAlaSerLeuGlySerArgAlaSerLeuSer	86
Db	2	CTGGCCCTGCTTGGCTCTCTGCTGCTGGTGGTACACCTGGGAGCTGGGCAACGCTTGCT	61
QY	87	AlaGlnGluProAlaGlnGluGluLeuValAlaGlnGluAspGlnAspProSerGluLeu	106
Db	62	GCCCGAGGAGCCCTTCACAGAGAGCTGACAGACAGACCGCGCGAGCCCTCGAATCG	121
QY	107	AsnProGlnThrGluGluSerGlnAspProAlaProPheLeuAsnArgLeuValArgPro	126
Db	122	AATCCCGACAGACAGAAAGCCAGCATGTGTCATCTTCTTGGAAACATGATGCGGCT	181
QY	127	ArgArgSerAlaProLysGlyArgLysThrArgAlaArgArgAlaLeuAlaHisTyr	146
Db	182	CGAAGAAGTGCCTCTAAAGGCCGGAAGCGCGGCT-CCCGAGCATATTCAGACCATATAT	240
QY	147	GluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlnAlaValAspGlyThrValSer	166
Db	241	GAGGTTCATCCTCGGGCCACGAGCATGGAGACAGACAGAGTGGATGGGACATGAGT	300
QY	167	GlyTyrPglGluAlaArgIleAsnSerSerProLeuArgTyrAsnArgGlnIleGly	186
Db	301	GGCTGGGAAGAACCAAAATCAACAGCTCCAGCCCTTGCCTACGACCGCGAGATTGGG	360
QY	187	GluPheIleValThrArgAlaGlyLeuLysTyrIleLeuTyrCysGlnValHisPheAspGlu	206
Db	361	GAAATTACAGTCATCAGGGCTGGGCTGTACTACCTGTACTGTACAGTGCACATTGATAG	420
QY	207	GlyLysAlaValTyrLeuLysLysAspLeuValAspGlyValLeuAlaLeuArgCys	226
Db	421	GGAAAGCGCTGTACTACTAGAGCTGGACATTGCTGGTGAACGGTGTCTGGCCCTGGCTGC	480
QY	227	LeuGlnGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArgLeuCysGln	246
Db	481	CTGGAAGAATTCTCAGCCACACACCAAGCTCTCCTGGGGCCCCACCTCGTTTGGCCAG	540
QY	247	ValSerGlyLeuLeuAlaLeuArgProGlySerSerIleArgIleArgThrLeuProThr	266
Db	541	GTGTCTGGGCTGTCCCTGCGGCGAGGTCT-TCCTCTTGGATCCGACCCCTCCCTGG	599
QY	267	AlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGlnValHis	284
Db	600	GCTCATCTTAAGGCTGC-CCCTTCTTACTACTTGTGAGCTCTTCAAGTTTAC	652

RESULT 7	948 bp	mRNA	linear	EST 16-JUL-2002
LOCUS	BO707185			
DEFINITION	AGENCOURT 8353983 NIH_MGC_113 Homo sapiens cDNA clone IMAGE:6278608			
ACCESSION	BO707185			
VERSION	BO707185.1			
KEYWORDS	EST.			
SOURCE	human.			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			

REFERENCE 1 (bases 1 to 948)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. Mark Watson
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LCM2466 row: n column: 17
High quality sequence start: 24
High quality sequence stop: 550.
Location/Qualifiers
1..948
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NIH-MGC_113"
/lab_host="DH10B (phage-resistant)"
/note="Organ: spleen; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH-MGC Library."
BASE COUNT 173 a 334 c 252 g 187 t 2 others
ORIGIN
Alignment Scores:
Pred. No.: 2,95e-76 Length: 948
Score: 909.50 Matches: 179
Percent Similarity: 98.35% Conservative: 0
Best Local Similarity: 98.35% Mismatches: 2
Query Match: 62.98% Indels: 2
DB: 14 Gaps: 1
US-09-245-198a-4 (1-284) x BQ707185 (1-948)
QY 103 ProSerGluLeuAsnProGlnThrGluSerGlnAspProAlaProPheLeuAsnArg 122
|||||
Db 3 CCGTCGGAATCGAATCCCGACAGAAAGAGCAGAGATCTCGCTTCGTAACCGA 62
QY 123 LeuValArgProArgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIle 142
|||||
Db 63 CTAGTTCGGCTCGAAGAACTGCACCTAAAGCGCGGAAACACGCGCTCGAAGCGCATC 122
QY 143 AAlAlAlHisTyrGluValHisProArgProGlyGlnAspGlyAlaGlnAlaGlyValAsp 162
|||||
Db 123 GCAGGCCATTATTAAGTTCTCCACGACCTGGACAGAGCAGACC---GCAGGIGTGAC 178
QY 163 GlyThrValSerGlyTyrGluGluAlaArgIleAsnSerSerSerProLeuArgTyrAsn 182
179 GGGACAGTAGTGCTGGGAGGAGGAGCAACAGCAAGCTCCAGCCCTCGCCTCAACAC 238
QY 183 ArgGlnIleGlyLupheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCysGlnVal 202
|||||
Db 239 CGCCACATCCGGAGATTATAGTTCACCCGGCTGGCTCTACTACTGTCTAGGTG 298
QY 203 HisPheAspGluLysAlaValTyrLeuLysLeuAspLeuValAspGlyValLeu 222
299 CACTTTGATGAGGGAGGAGGCTGTCTACCTGAACCTGACTTGTGTGGATGTGCTG 358
QY 223 AlaLeuArgCysLeuGluLupheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeu 242
Db 359 GCCCTCGCGCTGGTGAAGAAATTCACGACACCTGCGGCCAGTTCCCTCGGGCCCCAC 418
QY 243 ArgLeuCysGlnValSerGlyLeuLeuAlaLeuArgProGlySerSerLeuArgIleArg 262

|||||
Db 419 CGCCTCGCAGGCTGCTGGCTGGCCCTGCGCCAGGCTCCCTGCGAGCCGC 478
QY 263 ThrLeuProThrAlaHisLeuLysAlaAlaProPheLeuThrTyrPheGlyLeuPheGln 282
Db 479 ACCCTCCCTCGGCGCATCTCAAGGCTGCCCTTCTCTACACTTTCGACTTTCAG 538
QY 283 ValHis 284
|||||
Db 539 GTTCAC 544
RESULT 8
BO884231
LOCUS BO884231 940 bp mRNA linear EST 16-AUG-2002
DEFINITION AGENCOURT 8682031 lupski.scfatic.nerve Homo sapiens cDNA clone
IMAGE:6197488 5', mRNA sequence.
ACCESSION BO884231 GI:22276239
VERSION BO884231.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 940)
AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Dr. James R. Lupski
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLM13607 row: j column: 17
High quality sequence stop: 453.
Location/Qualifiers
1..940
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Lupski.scfatic.nerve"
/sex="male"
/tissue_type="sciatic nerve"
/dev_stage="adult, 70 yr"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6 (Life Technologies); Site_1: NotI; Site_2: SalI; cDNA made by oligo-dT priming. Directionally cloned using the following adaptors: 5'-TGCACCGACGCGTCCG-3' and 5'-TGCAGTGTCTAGATCGGAGCGGCGCCCT(15)-3'. Size selected > 1 kb for average insert length 1.87 kb. This is a primary library, non-amplified. Library constructed by Life Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor College of Medicine) and is available through Life Technologies."
BASE COUNT 170 a 325 c 255 g 190 t
ORIGIN
Alignment Scores:
Pred. No.: 9.73e-76 Length: 940
Score: 904.00 Matches: 183
Percent Similarity: 97.87% Conservative: 1
Best Local Similarity: 97.34% Mismatches: 0
Query Match: 62.60% Indels: 4
DB: 14 Gaps: 0
US-09-245-198a-4 (1-284) x BO884231 (1-940)
QY 101 GlnAspProSerGluLeuAsnProGlnThrGluSerGlnAspProAlaProPheLeu 120
|||||

Db 1 CAGGACCCGTCGAGACTGATATCCAGACAGAGAAAGCCAGATCTGCGCTTCTCTG 60
Qy 121 AsnArgLeuValArgProArgArgSerAlaProLysGlyArgGlyThrArgAlaArgArg 140
Db 61 AACCGACAGTGTGGCTCGCAGAGTGCACCTAAAGCCGGAACACAGCGGCTCGAAGA 120
Qy 141 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyValAlaGlnAlaGly 160
Db 121 GCGATCGAGCCGCAATTAAGAGTTATCATCCAGACCTGGACAGACGAGCGGACAGCT 180
Qy 161 ValAspGlyThrValSerGlyTyrPglGluAlaArgGlnLeuSerSerProLeuArg 180
Db 181 GTGACGCGGACAGTGAAGTGGCTGGAGGAGCAATCAACAGCTCCAGCCCTCTGGCC 240
Qy 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200
Db 241 TACACACCCGCGATCGGAGAGTTATAGTACACCGGCTGGGCTCTACTACTGACTGT 300
Qy 201 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220
Db 301 CAGGTGCACCTTGTAGAGAGGAGGAGGCTGTCTACCTGAAGCTGACTGCTGTGATGGT 360
Qy 221 ValLeuAlaLeuArgCysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyPro 240
Db 361 GTGCTGGCCCTGCGCTGCTGGAGGAATTCAGCCACTGCGGCGCAGTTCCTCGGCGCC 420
Qy 241 GlnLeuArgLeuCys-GlnValSerGlyLeuLeuAlaLeuArgProGlySer-SerLeuA 260
Db 421 CAGCTCCGCTCTGCGCCAGGTGTCTGGGCTGTGGCCCTGGCGGAGGAGGCTCTCGTC 480
Qy 260 TgIleArgThrLeuProThrAlaHisLeuLysAlaAla-ProPheLeuThrTyr-PheG 279
Db 481 GGATCGGACACCTCCCTCGGCGGCAATCTCAAGGCTGCCCTTCTCTCACTCTTGG 540
Qy 279 yLeuPheGlnValHis 284
Db 541 ACTCTTCAGAGTTTAC 556

RESULT 9
AM763237 561 bp mRNA linear EST 04-MAY-2000
LOCUS ur70d09.y1 NCI-CGAP_Mam3 Mus musculus cDNA clone IMAGE:3155633 5'
DEFINITION similar to YR:054907 054907 TNF-RELATED WEAK INDUCER OF APOPTOSIS
; mRNA sequence.
ACCESSION AM763237
VERSION AM763237.1 GI:7695174
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 561)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
OTHER ESTS: ur70d09.x1
CONTACT: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Lothar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMNL)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LMNL at:
<http://imgl.gov/image/html/lresources.shtml>
MGI:1058389
Seq primer: -40RP from Gibco
High quality sequence steps 433.
Location/Qualifiers
1..561
/organism="Mus musculus"

/strain="129, C57BL/6J, FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:3155633"
/clone_lib="NCI-CGAP_Mam3"
/tissue_type="tumor, gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
Site:2: NotI. Cloned unidirectionally. Primer: oligo dr.
library constructed by Life Technologies. Investigators
providing samples: Lothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."

BASE COUNT 108 a 158 c 194 g 100 t 1 others

ORIGIN

Alignment Scores:
Pred. No.: 1.52e-68 Length: 561
Score: 825.00 Matches: 161
Percent Similarity: 90.91% Conservative: 9
Best Local Similarity: 86.10% Mismatches: 17
Query Match: 57.13% Indels: 0
Db: 10 Gaps: 0

US-09-245-198A-4 (1-284) x AM763237 (1-561)

Qy 41 SerGlnArgArgArgGlyValArgGlyGluProGlyThrAlaLeuLeuValProLeuAla 60
Db 1 AGCCAGAGGCGGAGGGGCGCGGGGAGAGCGGGACCGCCCTCTGCGCCGCTGGCTG 60
Qy 61 LeuGlyLeuGlyLeuAlaLeuAlaCysLeuGlyLeuLeuLeuAlaValValSerLeuGly 80
Db 61 CTGACCTCGGGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTG 120
Qy 81 SerArgAlaSerLeuSerAlaGlnGluProAlaGlnGluGluLeuValAlaGluGluAsp 100
Db 121 AGCTGGGCAAGCGTGTGCGCCAGAGACCTTCTCAGAGAGAGCTGACAGAGAGAGCCG 180
Qy 101 GlnAspProSerGlyLeuAsnProGlnThrGluGluSerGlnAspProAlaProPheLeu 120
Db 181 CGGAGGCCCTCGAAGCTGATCCCGACAGAGAGAAACCCAGAGTGTGACTTCTTCTG 240
Qy 121 AsnArgLeuValArgProArgArgSerAlaProLysGlyArgGlyThrArgAlaArgArg 140
Db 241 GAACACAGTACTGGCGCTCGAAGAGTGTCTCTAAAGCGGAGCGGCGCTGCGCGA 300
Qy 141 AlaIleAlaAlaHisTyrGluValHisProArgProGlyGlnAspGlyValAlaGlnAlaGly 160
Db 301 GCTATTGACGCCAATTAAGAGTTATCATCCAGACCTGGGCTCTACTACTGACTGT 360
Qy 161 ValAspGlyThrValSerGlyTyrPglGluAlaArgGlnLeuSerSerProLeuArg 180
Db 361 GTGATGGGACAGTGAAGTGGGAGAGGACCAAAATCAACAGCTCCAGGCTTGGAGC 420
Qy 181 TyrAsnArgGlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrTyrLeuTyrCys 200
Db 421 TACAGACCGCGAATTAAGAGTTATCATCCAGACCTGGGCTCTACTACTGACTGT 480
Qy 201 GlnValHisPheAspGluGlyLysAlaValTyrLeuLysLeuAspLeuValAspGly 220
Db 481 CAGGTGCACCTTGTAGAGAGGAGGAGGCTGTCTACCTGAAGCTGACTGCTGTGAACG 540
Qy 221 ValLeuAlaLeuArgCysLeu 227
Db 541 GTGCTGGCCCTGCGCTGCTG 561

RESULT 10
B0671259 963 bp mRNA linear EST 15-JUN-2002
LOCUS B0671259
DEFINITION AGENCOIRT_8303564 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6274716
5', mRNA sequence.
ACCESSION B0671259
VERSION B0671259.1 GI:21782093

Oy 138 AAlArGArGaLaIleAlAAlHISTYGLVAlHisProArgProGLyGlnAspGLyAla 157

Db 12 GCCTGAAGACCGATCCAGCCCATTTATGAAGTTCATCCACGACCCTGGACACGACGAGC 71

QY	158	G1NA1AG1VAlAASg1YThrValSerclYtrpg1ug1uAlAATg11AASnSerSerSer	177
QY	158	G1NA1AG1VAlAASg1YThrValSerclYtrpg1ug1uAlAATg11AASnSerSerSer	177
Db	72	CAGCGAGGTGTGACAGGGACAGTGTGTGTGGAGGAAGCAGATCAACAGCTCCAGC	131
QY	178	ProLeuAaGTYrAsnArG1n11eG1yS1uPhel1eVal1ThrArGAlAG1YLeuTYrYr	197
Db	132	CCTCGGCCCTACAAACCGCAGATCGGGAGATTATAGTACACCCGGGCTGGGCTACTAC	191
QY	198	LeuTYrCYsG1nVal1HisPheAspG1uG1yS1aVal1TYrLeuLYsLeuAspLeuLeu	217
Db	192	CTGTACTCTCAGAGTGCACATTGATAGGGGAAGGCTGTCTACTGAACCTGACTTGGCTG	251
QY	218	ValAspG1VAl1LeuAlA1LeuA1ArgYs1LeuG1uG1uPheserAlaThrAlA1AsSerSer	237
Db	252	G1GAGTGTGTGTGTGGCCCTGTGGCTGTGGAGGAATTCACGCCATCGCGCGAGTTCC	311
QY	238	LeuG1YrProG1nLeuA1ArgYsG1nVal1SerG1yLeuLeuAlA1LeuA1ArgProG1ySer	257
Db	312	CTCGGGGCCCAAGCTCGGCTCTGCCAGATGTCTGTGGCTGTGGCCCTTGGCCCTCGGCGAAGTCC	371
QY	258	SerLeuA1Arg11eATgThrLeuProTrpAlA1HisLeuLYs1aValA1aProPheLeuThrTYr	277
Db	372	TCCCGCCGAGTCCGACCCCTCCCTGGGCGCATCTCAAGAGCTGCCCTTCTCTACCTAC	431
QY	278	PheG1yLeuPheG1nVal1His 284	
Db	432	TTCGGACCTTTCACAGTTTCAC 452	
RESULT 12			
LOCUS	B0674188	951 bp	mRNA linear EST 15-JUL-2002
DEFINITION	AGENCOURT_8354100 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6275564		
ACCESSION	B0674188		
VERSION	B0674188.1	GI:21785022	
KEYWORDS	EST.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
AUTHORS	NIH-MGC http://mgc.ncl.nih.gov/.		
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)		
JOURNAL	Unpublished (1999)		
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cga@b.jrnl.nih.gov Tissue Procurement: ATCC cDNA Library Preparation: Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA sequencing by: Agencourt Bioscience Corporation cDNA distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov Plate: L10M2459 row: d column: 01 High quality sequence stop: 335.		
FEATURES	Location/Qualifiers		
SOURCE	1..951		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:6275564"		
	/clone_id="NIH_MGC_102"		
	/tissue_type="epidermoid carcinoma, cell line"		
	/lab_host="DH10B (phage-resistant)"		
	/note="Organ: salivary gland; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCAGCAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."		
BASE COUNT	165 a 286 c 315 g 184 t 1 others		
ORIGIN			

Alignment Scores:				
Pred. No.:	2,356-61	Length:	951	
Score:	753.00	Matches:	152	
Percent Similarity:	95.62%	Conservative:	1	
Best local Similarity:	95.00%	Mismatches:	5	
Query Match:	52.15%	Indels:	2	
DB:	14	Gaps:	0	

US-09-245-198A-4 (1-284) x B0674188 (1-951)				
QY 124	valaAProAtgArgSerAlaProLysGlyArgLysThrArgAlaArgAlaIleAla	143		
Db 2	GTTCGGCCCTCGCAGAAATGACACTAAAGCCCGGAAACACGGGGCTCGAAGACGATCGCA	61		
QY 144	AlaHs1YrGluValHs1ProArgrProGlyGlnAspGlyAlaGlnAlaGlyValAspGly	163		
Db 62	GCCCAATTATGAAGTTCATCCACGACCTCGACAGGAGCGGACGCGAGGTGTGACGGG	121		
QY 164	ThrAlaSerGlyTrpGluGluAlaArgIleAsnSerSerProLeuArgrTyraArgr	183		
Db 122	ACAGTGAATGGCTGGGAGGAAACCGAATCAACACTCCAGCCCTCGCGCTACAAACCGC	181		
QY 184	GlnIleGlyGluPheIleValThrArgAlaGlyLeuTyrrTyrrLeuTyrrCysGlnValHs	203		
Db 182	CAGATCGGGAGATTATATGATCATCCCGGGGCTGGCTCTACTACCTGATGTCAGGTGCAC	241		
QY 204	PheAspGluGlyLysAlaValTyrrLeuLysLeuAspLeuValAspGlyValLeuAla	223		
Db 242	TTTGATGAGGGGAAAGGCTGTCTACTCGAAGCTGAGCTGTGTGTGATGTGTGCTGGCC	301		
QY 224	LeuArGysLeuGluGluPheSerAlaThrAlaAlaSerSerLeuGlyProGlnLeuArg	243		
Db 302	CTGGCCTGCTCGAGAAATCTCAGCCACCTGGGGCAGTTCCTCTGGGGGCCACGCTCCGC	361		
QY 244	LeuGysGlnValSerGlyLeuLeuAlaLeuArg-ProGlySerSerLeuArgIleArgTh	263		
Db 362	CTCTCCAGAGTGTGTGGGCTGTGGCCCTCGGGGCGAGGGTCTCTCCGTGGGAGTCCCAAC	421		
QY 263	IleuProThrAlaHs1LeuLysAlaAla-ProPheLeuThrTyrrPheGlyLeuPhe	281		
Db 422	CCTCCCGGGGGCCATTAAAGGGTGGCCCTCTCTCACCCTATTTTGGAAATTTT	477		

RESULT 13				
LOCUS B1762908	785 bp	mRNA	linear	EST 25-SEP-2001
DEFINITION 603047966F1 NIH_MGC_116 Homo sapiens cDNA clone IMAGE:5188168 5',				
ACCESSION B1762908				
VERSION B1762908.1	GI:15754486			
KEYWORDS EST.				
SOURCE human.				
ORGANISM Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.				
1 (bases 1 to 785)				
NIH-MGC http://mgc.nci.nih.gov/				
National Institutes of Health, Mammalian Gene Collection (MGC)				
Unpublished (1999)				
Contact: Robert Strausberg, Ph.D.				
Email: cga@bts-remail.nih.gov				
Tissue Procurement: Life Technologies, Inc.				
cDNA Library Preparation: Life Technologies, Inc.				
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LMLN)				
DNA Sequencing by: Incyte Genomics, Inc.				
Clone distribution: MGC clone distribution information can be				
found through the I.M.A.G.E. Consortium/LMLN at:				
http://image.lnl.gov				
Plate: L1AM11470	row: C	column: 17		
High quality sequence stop: 647.				
Location/Qualifiers				
1. 785				

FEATURES	source
source	/organism="Homo sapiens"

Db	137	GGGAGCCGGGACCGCCCTCGTGCCTCGGCGTGGGCGCTGGGCGCTGCTGCTGC	196
Oy	68	acgyleuglyleuenteuleuAlaValSerLeuGlySerArgAlaSerLeuSerAlaGl	88
Db	197	CTGGCTGGGCTCTCTGCTGGCGCGGTGCTAGTTTGGGGCGCGGCGATGCTGTCGCCCA	256
Oy	88	nglprAlaaglIngluIuLeuValAlaGluGlnAspAlnAspAlnAspProSerGluLeuAsnPr	108
Db	257	GGACCTGCCAGAGAGAGAGCTGTGGCGAGAGAGAGACCAGGACCCCTCGGAAGTAATGC	316
Oy	108	ogInThrGluGlnuSerGlnAspProAlaPheProLeuAsnArgLeuValArgProArgAr	128
Db	317	CCACACGAAAGAAAGCCAGATCTGCGCCCTTTCTTGAAACACACTAGTTTGGCCCTCGCAG	376
Oy	128	gSerAlProLysGlyArgLysThrArgAlaArgArgAlaIleAlaIleAsnIleAsnIleVala	148
Db	377	AAGTGCACTTAAGCGCGGAAACACGCGGCTCGAAGACGATTCGACGCCCATATGAAGT	436
Oy	148	lHisProArgProGluGlnAspGlyAlaGlnAlaGlyValaAspGly	163
Db	437	TCATCCACGACCTGACAGCAGCAGCGCCAGCA-----GATGGA	476
RESULT 15			
AKO20909			
LOCUS	AKO20909	1033 bp	mRNA
DEFINITION	Mus musculus adult retina cDNA, RIKEN full-length enriched library, clone:A930030D13:tumor necrosis factor (11gag) superfamily, member 12, full insert sequence.		HTC 19-JAN-2002
ACCESSION	AKO20909	GI:12861640	
KEYWORDS	HTC; CAP trapper.		
SOURCE	Mus musculus (Strain:C57BL/6J) adult retina cDNA to mRNA, clone:11b:RIKEN full-length enriched mouse cDNA library clone:A930030D13.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.		
AUTHORS	Carninci,P. and Hayashizaki,Y.		
TITLE	High-efficiency full-length cDNA cloning		
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)		
MEDLINE	99279253		
PUBMED	10349636		
REFERENCE			
AUTHORS	2 Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komoto,H., Okazaki,Y., Muramatsu,M., and Hayashizaki,Y.		
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes		
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)		
MEDLINE	20499374		
PUBMED	11042159		
REFERENCE			
AUTHORS	3 Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., Komoto,H., Akiyama,D., Nishi,K., Kitsuana,T., Tashiro,H., Itoh,M., Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A., Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K., Fujisake,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watabiki,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura,S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kita,A. and Hayashizaki,Y.		
TITLE	RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer		
JOURNAL	Genome Res. 10 (11), 1757-1771 (2000)		
MEDLINE	20530913		
PUBMED	11076861		
REFERENCE			
AUTHORS	4 Kawai,J., Shinagawa,A., Shibata,K., Yoshino,M., Itoh,M., Ishii,Y., Aikawa,T., Hara,A., Fukunishi,Y., Komoto,H., Adachi,J., Fukuda,S., Aizawa,K., Nishi,K., Kiyosawa,H., Komodo,S., Yamanaka,I., Saito,T., Okazaki,Y., Gotohori,T., Bono,H., Kasukawa,T., Saito,R., Kadoya,K., Matsuda,H., Ashburner,M., Batalov,S., Casavant,T., Kleeschmann,W., Gaasterland,T., Gissi,C., King,B., Kochiwa,H., Knehl,P., Lewis,S., Matsuno,Y., Nikaido,I., Pesole,G.,		

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Quackenbush, J., Schriml, L.M., Staab, F.F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J.R., Boffelli, D., Boujanga, N., Carninci, P., de Bonaldo, M.F., Brownstein, M.J., Bolt, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, M., Hunne, D.A., Kamiya, M., Lee, N.H., Lyons, P., Marchionni, L., Mashima, J., Mazzarello, J., Mommaerts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seya, T., Shibata, Y., Storch, K.F., Suzuki, H., Toyooka, K., Wang, K.H., Weitz, C., Whitteker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohlski, S. and Hayashizaki, Y.

TITLE
Functional annotation of a full-length mouse cDNA collection

Nature 409 (6821), 685-690 (2001)

21085660
11217851

5 (bases 1 to 1033)

Akachi, J., Alzawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Atakawa, T., Baldarelli, R., Bonin, H., Brownstein, M., Bolt, C., Carninci, P., Fukuda, S., Fukushima, Y., Furuno, M., Hangaki, T., Harada, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Horii, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numasaki, R., Ono, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Salto, H., Satou, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shinagawa, A., Shiraki, T., Soegabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takashi, F., Tanaka, T., Teijima, Y., Toyo, T., Yamamura, T., Yamanaka, I., Yanunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE
JOURNAL
DIRECT SUBMISSION

Submitted (18-AUG-2000) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration and Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suenho-cho, Tsutsumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@sc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

Please visit our web site (<http://genome.gsc.riken.go.jp/>) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'-GAGAGAACAACATCCAAAGCTCTTTTTCCTTTTTTTTGA 3'] cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through two rounds of normalization to Rot = 20.0 and subtraction to Rot = 458.8. Second strand cDNA was prepared with the primer adapter of sequence [5'-GAGAGAACAATTTCGACTTAATTATATCCTCCCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmid vector KS(+) after bulk excision from Lambda phage I. Cloning sites: 5' end: SalI; 3' end: BamHI. Host: DH10B. Retina RNA was provided by Stefano Gustincich (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged.

FEATURES
SOURCE

1..1033

/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="FANTOM_DB:A930030D13"
/db_xref="MGD:MGI:1911867"
/db_xref="taxon:10090"
/clone="A930030D13"
/tissue.type="retina"
/clone.lib="RIKEN full-length enriched mouse cDNA library /dev_stage="adult"

RS 5/15/03

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 05:28:59 ; Search time 54.1218 Seconds
(without alignments)
1081.216 Million cell updates/sec

Title: US-09-245-198A-4
Perfect score: 1444
Sequence: 1 MSLLDFRISARRLPLPRSLG.....PMAHLKAPLTYGFLQVH 284

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: SPREMBL_21:*
2: sp_bactera:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_plant:*
10: sp_rodent:*
11: sp_virus:*
12: sp_vertebrate:*
13: sp_unclassified:*
14: sp_rvivirus:*
15: sp_bacteriap:*
16: sp_archaeap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	116	8.0	325	5	Q9V5G2 drosophila
2	103	7.1	565	16	Q9KY66 streptomyce
3	101	7.0	643	16	Q9KZ17 streptomyce
4	100.5	7.0	375	16	Q9RRH5 streptomyce
5	100	6.9	850	11	Q9J115 mus musculus
6	99.5	6.9	378	16	Q9AAB9 caulobacter
7	99	6.9	614	4	Q8WU14 homo sapien
8	99	6.9	855	4	Q9NYK9 homo sapien
9	99	6.9	878	4	Q9UFU7 homo sapien
10	98	6.8	206	16	Q9S2W5 streptomyce
11	97.5	6.8	1560	4	Q9GJP2 homo sapien
12	97	6.7	408	10	Q8S515 oryza sativ
13	97	6.7	926	4	Q9NRA0 homo sapien
14	97	6.7	1058	4	Q9Y4G2 homo sapien
15	96	6.6	629	10	Q8S1A6 oryza sativ
16	95.5	6.6	536	4	Q9HB96 homo sapien

17	95.5	6.6	655	16	Q9FBR7 streptomyce
18	94.5	6.5	748	5	Q8T2Y0 trypanosoma
19	94.5	6.5	1696	11	Q9WTR8 ratius norv
20	94	6.5	937	16	Q9JUD1 streptomyce
21	93.5	6.5	776	16	Q9RJ01 streptomyce
22	93.5	6.5	854	16	Q9FZP0 streptomyce
23	93.5	6.5	1038	10	Q9AS09 oryza sativ
24	93	6.4	507	2	Q9RMT5 streptomyce
25	93	6.4	1043	10	Q93VC8 oryza sativ
26	92.5	6.4	484	11	Q9CTA7 mus musculus
27	92.5	6.4	614	4	Q9NW41 homo sapien
28	92.5	6.4	707	11	Q9Z5J8 mus musculus
29	92.5	6.4	1511	13	Q91BE2 fugu rubrip
30	92	6.4	351	12	Q8S302 orf virus
31	92	6.4	566	16	Q70005 streptomyce
32	92	6.4	845	4	Q96H68 homo sapien
33	92	6.4	977	4	Q8TE71 homo sapien
34	91.5	6.3	212	16	Q86312 mycobacteri
35	91.5	6.3	221	10	Q8RYZ6 oryza sativ
36	91.5	6.3	664	16	Q9HYT9 pseudomonas
37	91.5	6.3	3169	5	Q917V8 drosophila
38	91	6.3	681	16	Q9HUZ6 pseudomonas
39	91	6.3	1217	4	Q60336 homo sapien
40	90.5	6.3	449	2	Q93HD5 streptomyce
41	90.5	6.3	760	11	Q8VD14 mus musculus
42	90.5	6.3	1100	2	Q937L3 delinococcus
43	90.5	6.3	1329	16	Q9CD30 mycobacteri
44	90	6.2	285	2	Q33629 saccharopol
45	90	6.2	422	16	Q910B9 streptomyce

ALIGNMENTS

RESULT 1

ID	Q9V5G2	PRELIMINARY;	PRT;	325 AA.
AC	Q9V5G2:			
DT	01-MAY-2000 (TREMBLrel. 13, Created)			
DT	01-MAY-2000 (TREMBLrel. 13, Last sequence update)			
DT	01-JUN-2001 (TREMBLrel. 17, Last annotation update)			
DE	CG12919 protein.			
GN	CG12919.			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_TaxID=7227;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BERKELEY;			
RX	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,			
RA	George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,			
RA	Sutcliffe G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Heit G., Nelson C.R., Miklos G.L.G.,			
RA	Abriil J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,			
RA	Ballem R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Besson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brodeur P., Brothier P.,			
RA	Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	de Pablos K., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,			
RA	Dubin K.J., Evangelista C.C., Ferreira C., Ferreira S., Fleischmann W.,			
RA	Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,			
RA	Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,			
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jajali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merulov G., Mishina N.V., Moberly C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazolo M., Piltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Relbert C., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spler E., Spredling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitsks R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Nassarman D.A., Weinstein G.M., Weisenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of *Drosophila melanogaster*.";
 RT Science 287:2185-2195(2000).
 DR EMBL: AE003831; AAF58848.1; -;
 DR Flybase: FBgn0033483; CG12919.
 DR InterPro: IPR000478; TNF_family.
 DR SMART: SM00207; TNF_1.
 DR PROSITE: PS00251; TNF_1; 1.
 DR PROSITE: PS50049; TNF_2; 1.
 SQ SEQUENCE 325 AA; 36862 MW; 6E5CBB6964F1A3A CRC64;

Query Match 8.0%; Score 116; DB 5; Length 325;
 Best Local Similarity 25.4%; Pred. No. 0.07;
 Matches 57; Conservative 32; Mismatches 97; Indels 38; Gaps 10;

OY 82 RASLSQAEFAQELVAE-----EPDPSSELNPTQESQDPAFLNR-----LYRPR 127
 DB 118 RRSRSIADVRNEQNTQGHTELOEKSSNEATSKES--PAPLHRRRMRSHHLLVLRKG 175
 OY 128 RSAPKRRKTRARAIAAHVEHPRPQDGAQAGVDGTVSGMEARLNSSPLRYNPOIGE 187
 DB 176 ESLLSARSDSR--AAHPLLSRRHGGSM--CYHCDMTYGNMNNENSTYOG--HFGRRDGV 231
 OY 188 FIVTRAGLYLYCOV-----HEDEKAVYLYKLDDLVDGVALRCLCEFSATASSLGPOI 242
 DB 232 LFTVTNGLYVYVQAQICYNNSHDNGFIVF-----OGDTPFLOCLN---TVPNMPHKV 281
 OY 243 RLCQVSGLLALRPGSSLRIRTL--PWAHLKAPFLTYRGLPOV 283
 DB 282 HHCHTSGLIHLERNERIHLDIHDRNAVALREGNNRSTFEIRKV 325

RESULT 2
 O9KY66 PRELIMINARY: PRT; 565 AA.

AC O9KY66: 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative peptidase.
 GN SC04798 OR SC063A.09C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID-1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Brown S.P., Harris D.;
 RL Submitted (May-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RL Submitted (May-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2);
 RA Thompson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RX MEDLINE-97000351; PubMed-8843436;

RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,
 RA Kinschi H., Hopwood D.A.;
 RT "A set of ordered cosmids and a detailed genetic and physical map for
 the 8 Mb streptomyces coelicolor A3(2) chromosome.";
 RT Mol. Microbiol. 21:77-96(1996).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RA Bentley S.D., Chater K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
 RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
 RA Huang C.-H., Kleser T., Larke L., Murphy L., Oliver K., O'Neill S.,
 RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
 RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
 RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
 RA Hopwood D.A.;
 RT "Complete genome sequence of the model actinomycete Streptomyces
 coelicolor A3(2)."
 RT Nature 417:141-147(2002).
 DR EMBL: AL356832; CAB92661.1; -;
 DR InterPro: IPR002886; Peptidase_M37.
 DR Pfam: PF01551; Peptidase_M37; 1.
 SQ SEQUENCE 565 AA; 58070 MW; 7D0418D480C6A284 CRC64;

Query Match 7.1%; Score 103; DB 16; Length 565;
 Best Local Similarity 21.2%; Pred. No. 1.7;
 Matches 70; Conservative 33; Mismatches 99; Indels 128; Gaps 16;

OY 2 SLIDFELSA---RLPLRSLGSDGAVRQAPPMARARSGRRGRGEGTALL-- 56
 DB 290 ALLDDEEATPDARIPVARR--GSRAARSRRQP-----AKRS-----ALLTI 331
 OY 57 -VPLALGL--GLALACILLLAVVSLGSRASLSAQEPAQEE----- 94
 DB 332 AVPSACVMSVAGIAAASVGLTG--DEGTETASADPDGAEAPVPPSAANKKLDTOIL 389
 OY 95 -LVAED-----QDPSELNPTQESQDPAFLNRVPRNSAPKGRTRARRAIAHY 146
 DB 390 SLAAGADDFADRASRTOERIDLKAEDDAEKRAAQEARKERLRPFALPVKQHGSAVY 449
 OY 147 EVNRPQDGS-----AQAGVDGTVSGMEARLNSSPLRYNPOIGE 187
 DB 450 -----GQAGINMSSHTGIDFPVLOGTIVMAKTDGVR-----TQFNSAYGN 491
 OY 188 FIVTRA--GLYLYCOVH-----FDEGKAVYLYKLDDLVDGVALRCLCEFSATAAS 236
 DB 492 MMIVTAKDGTETWYCHLSSQVPSGTYVAKGDAI-----AYSGDSGN 533
 OY 237 SLGPQLRLCOVSGLLALRPGSSLRIRTLPW 266
 DB 534 STGPHLH-----FVVRPAGSSIDPLPW 556

RESULT 3
 O9KZ17 PRELIMINARY: PRT; 643 AA.

AC O9KZ17: 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Hypothetical protein SC02220.
 GN SC02220 OR SC10B7.15.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID-1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-A3(2) / M145;
 RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
 RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kleser H.,
 RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,


```
Db 601 LGSPSTPPALNGNPPRSPPEPRRAGPYRENGSSGTPKLIORALRCGTA-----LLA 654
QY 76 VVSLG-----SR---ASLSAQEPAP-----BELVAEEDQDPELNPQTEESODPA 117
Db 655 SLGLGRDLQPPGSLGRSGSPPTAPPPAPSCPPPLPSTPLIRLSQTPPDHSSPTPG 714
QY 118 PFLNRL-----VPRPRSAFKGKTRARRALAHYEVHPRPGDGAQAQVGDCT 164
Db 715 PLLLDLGVPGSPSAKSPRRETRGR-----TVSPPEISRSAPETPCT 758

RESULT 6
Q9AAB9 PRELIMINARY; PRT; 378 AA.
ID 09AAB9;
AC 01-JUN-2001 (Tremblrel. 17, Created)
DT 01-JUN-2001 (Tremblrel. 17, Last sequence update)
DE 01-MAR-2002 (Tremblrel. 20, Last annotation update)
DE H1YD family secretion protein.
GN CC0683.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter;
OX NCBI_TaxID=155892;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE=21173698; Pubmed=11259647;
RA Nielsen W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Eلسen J., Heidelberg J.F., Alley M.R.K., Onta N., Madcock J.R.,
RA Pollock A., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Utecherback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005743; AAK22668.1;
DR TIGR: CC0683;
DR InterPro: IPR003997; RtxD.
DR PRINTS: PR01490; RTXTOXIND.
KW Complete proteome.
SQ SEQUENCE 378 AA; 39659 MW; 88B818C437C15CBC CRC64;

Query Match 6.9%; Score 99.5; DB 16; Length 378;
Best Local Similarity 25.1%; Pred. No. 2.1;
Matches 73; Conservative 29; Mismatches 106; Indels 83; Gaps 13;

QY 19 LGSRDGAVRQAQPPAPMAARSRQRGR--GEGTALLVPLATLGLALCLGLLLA 75
Db 36 LAAROGGIVANVH-----REGDRVVGQPIATLTLLSALRGDSFAVLSR--- 81
QY 76 VVSLGSAASLSAQEPAP--QBELVAEEDQDPELNPQTEESODPAFPLNRLVPRRS--AP 131
Db 82 --SLGAQSSAAGSRAATQAALAAESRQ-----LVQRRALNRELAEAKRSALQ 139
QY 132 KGRKTRAR-----RAIAAHYEVHPRPGDGAQAQVGDTSQGEAKRINSSPLRYRQI 185
Db 130 TERLTTLARAFAVARETTIAAGFLSPRE---LQARRSAELAVQDEASTLSQALSTYRQI 185
QY 186 GEFIVTRAGLYLYLYCOVHDEGRAVYLKLDL---LVYDGVIALRCLREFSATASSLGPOL 242
Db 186 GEV-----DARLAATIPIDQLAARAEAASTLGLGEG-QATQVYEAQGRYV 227
QY 243 RLCOVSGLLA-----LRPGSLRIRTIPLMAHLKAPPL 275
Db 228 VVATVAGRVAAALPVEAQGTVGAAVAVALTPGDSALVAEL-VAPSRAGFV 277

RESULT 7
Q8WU14 PRELIMINARY; PRT; 614 AA.
ID 08WU14
```

```
AC Q8WU14; Q9BR73;
DT 01-MAR-2002 (Tremblrel. 20, Created)
DT 01-MAR-2002 (Tremblrel. 20, Last sequence update)
DE 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Histone deacetylase 7A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-COLON;
RA Strausberg R.;
RN Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RC TISSUE-B-CELL;
RA Strausberg R.;
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC020505; AAN20505.1;
DR EMBL: BC006453; AAN06453.1;
DR InterPro: IPR000286; His_deacetylase.
DR Pfam: PF00850; Hist_deacetyl; 1.
DR PRINTS: PR01270; HDASUPER.
SQ SEQUENCE 614 AA; 66186 MW; 2B638D8C666B502 CRC64;

Query Match 6.9%; Score 99; DB 4; Length 614;
Best Local Similarity 24.2%; Pred. No. 4.2;
Matches 46; Conservative 21; Mismatches 67; Indels 56; Gaps 6;

QY 11 RLPLPRSLGSRGCAVR-----QAQPPAPMAAR-----RSQRRGR- 47
Db 63 RQIPSAEDLTEDDGCGQVVDLGLHRELHGHGPEARGAPALQOHPOVLLMEQORLAGRL 122
QY 48 -REPGTALLVPLATLGLALACTGLLAVVSLGSRASISAOEPAOEBLVAEEDQDPEL 106
Db 123 PRGTGTVLLPLAOGGHRPLS-----RAOSSPAAPASISAPPPASQARYLSSETPART 177
QY 107 NPQT-----EESODP-----APFLNRLVPRRSAPKGRKTRAR 139
Db 178 LPFTTGILYDSVLMKHQCGSCGDSRHPEHAGRIQSIWSRLQERGLMSGCCCLGRASLE 237
QY 140 RAIAAHYEVH 149
Db 238 ELQSVHSERH 247

RESULT 8
Q9NFK9 PRELIMINARY; PRT; 855 AA.
ID 09NFK9;
AC 09NFK9;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-JUN-2001 (Tremblrel. 17, Last annotation update)
DE Histone deacetylase 7.
GN HDAC7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-CERVICAL CARCINOMA;
RA Li S., Fischle W., Verdin E., Walsh M.J.;
RT "A novel class II HDAC is associated with the transcriptional
homeodomain repressor CCAAT displacement protein.";
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE239243; AAF63491.1;
DR InterPro: IPR000286; His_deacetylase.
DR Pfam: PF00850; Hist_deacetyl; 1.
DR PRINTS: PR01270; HDASUPER.
SQ SEQUENCE 855 AA; 92363 MW; 3157426AE1913DB6 CRC64;
```



```
OY 117 -----APFLNR--LVPRRSAPKGRKTRARRAIAAHYEVHPRPDGAGVDTYSG 167
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 337 YRGTPGSRPGLRHRFSQEPKRCISGLALDQ-----CVPSPGRQAQAPR---QG 384
OY 168 WEARL-----NSSSPLRYNRQIGEFIVTRAGLY-----YLICQVHFDEGKAVYLKLDL 217
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 385 HKSPFVYHRRQMGSLNPRGLMKLG--TVERRGAMGIMKELFCELSPLEFR-LYLSNE-- 439
OY 218 VDGVALRCLEEFSAATASSIGP-----QLRLCQVSGLLALRPGS 257
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 440 -----EHTCVENCSSLRCESVGPAPHSDFRFLVFSGKRLALRASS 479

RESULT 14
OY94G2 PRELIMINARY; PRT; 1058 AA.
AC OY94G2;
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE KIAA0356 protein (Fragment).
GN KIAA0356.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=BRAIN;
RX MEDLINE=97349984; PubMed=9205841;
  Nagase T., Ishikawa K., Nakajima D., Ohira M., Seki N., Miyajima N.,
  Tanaka A., Kotani H., Nomura N., Ohara O.;
  "Prediction of the coding sequences of unidentified human genes. VII.
  The complete sequences of 100 new cDNA clones from brain which can
  RT code for large proteins in vitro.";
  RL DNA Res. 4:141-150(1997).
DR EMBL; AB002354; BAA20813.2;
DR InterPro; IPR002219; DAG_PE-bind.
DR InterPro; IPR000532; Glucagon.
DR InterPro; IPR001849; PH.
DR InterPro; IPR004012; Run.
DR Pfam; PF00169; PH; 1.
DR PRINTS; PR00275; GLUCAGON.
DR SMART; SM00109; C1; 1.
DR SMART; SM00233; PH; 2.
DR PROSITE; PS50081; DAG_PE_BIND_DOM_2; 1.
DR PROSITE; PS50003; PH_DOMAIN; 1.
FT NON_TER 1 1
SQ SEQUENCE 1058 AA; 117671 MW; CD1AD1869C8C9D1 CRC64;

Query Match 6.7%; Score 97; DB 4; Length 1058;
Best Local Similarity 24.2%; Pred. No. 12;
Matches 69; Conservative 35; Mismatches 125; Indels 56; Gaps 13;

OY 4 LDFEISARRLPRLRSLSGRDGVAGVQAQPPAPMARRSQRRRGGEFTALVLAQL 63
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 332 LHCEPAPEPLPQAASGTQDVHVEPPRQAESPDLQOPEVSTSGQPPSSVSTAREV 411
OY 64 GLALACILGLAVVSLGSRASL---SAQEPQAEELVAEED--ODPSELNPTQTESODP--- 116
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 412 GCG---NGLQKAQAHGAGLGLVSSPTSPKKSWSIEDDFRPSREOPLESASDHPIAS 468
OY 117 -----APFLNR--LVPRRSAPKGRKTRARRAIAAHYEVHPRPDGAGVDTYSG 167
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 469 YRGTPGSRPGLRHRFSQEPKRCISGLALDQ-----CVPSPGRQAQAPR---QG 516
OY 168 WEARL-----NSSSPLRYNRQIGEFIVTRAGLY-----YLICQVHFDEGKAVYLKLDL 217
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 517 HKSPFVYHRRQMGSLNPRGLMKLG--TVERRGAMGIMKELFCELSPLEFR-LYLSNE-- 571
OY 218 VDGVALRCLEEFSAATASSIGP-----QLRLCQVSGLLALRPGS 257
```

```
DB 572 -----EHTCVENCSSLRCESVGPAPHSDFRFLVFSGKRLALRASS 611
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|

RESULT 15
OY81A6 PRELIMINARY; PRT; 629 AA.
AC OY81A6;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE B114B07.16 protein.
GN B114B07.16.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
  clone:B114B07.";
RL Submitted (FEB-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AP003334; BAB90006.1;
SQ SEQUENCE 629 AA; 68639 MW; 724C08A73DCB785D CRC64;

Query Match 6.6%; Score 96; DB 10; Length 629;
Best Local Similarity 20.8%; Pred. No. 7.7;
Matches 67; Conservative 33; Mismatches 116; Indels 106; Gaps 11;

OY 12 RLPFLRSLSRGGVAGVQAQPPAPMARRSQRRRGGEFTALVLAQL 55
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 107 RRPYRPEPLREAHVLRQAQPPRRRVRGRRERRRDRAVPDVEDDGDGSRSLRRRRRLS 166
OY 56 LVPALGLG-----LALACL----- 70
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 167 LADTAVAGLSGDVARRRHGLGCRSGSGGGGGGGLGGASHATPALTALIPRGTAFCRG 226
OY 71 -GLLIAVVSLSGRAS---LSAQEPQAEELVAEEDODPSELNPTQTESODPAPFLNRLVR 125
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 227 RGAALAPGSLGRRRCGCGRLAGVGAEREADDEGRDPGVEDRGRDGEPAEEVALYV- 285
OY 126 PRRSAPKGRKTRARRAIAAHYEVHPRPDGAGVDTYSGWEARLINSPLRYNRQI 185
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 286 ---ERGGGGRVYRRRRHHWTRERERERESAAGD-----RRRT 324
OY 186 GEFTVTRAGLYLYCQVHFDEGKAVYLKLDLVGVLAIRCLEEFSAATASSIGPOLRLC 245
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 325 LEYVAT--CVHGLVCPHIAATATRSL-----AACVLPYVAVMNARALRSIGPG-RQC 372
OY 246 OYSGLLALRPGSSLRIRLPMA 267
      |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!| |!|
DB 373 -VAG-----RCKSPYVPTRDWS 388
```

Search completed: March 31, 2003, 07:46:23
Job time : 62.1218 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: March 31, 2003, 07:44:21 ; Search time 28.4558 Seconds
(without alignments)
959.459 Million cell updates/sec

Title: US-09-245-198a-4

Perfect score: 1444

Sequence: 1 MSLLDFEISARRLPPLPSRLG.....PMAHLKAAPFLTYGRLQVH 284

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : PIR_73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
1	106.5	7.4	254	2	4-1BB ligand - hum
2	104	7.2	1323	2	N-methyl-D-asparta
3	100.5	7.0	375	2	hypothetical prote
4	99.5	6.9	378	2	HLyd family secret
5	99	6.9	878	2	hypothetical prote
6	98	6.8	206	2	probable membrane
7	97	6.7	441	2	mitosis-specific c
8	93.5	6.5	776	2	probable cation-tr
9	93.5	6.5	1657	2	hypothetical prote
10	93	6.4	310	2	hypothetical prote
11	92	6.4	351	2	ORF5 protein - Orf
12	92	6.4	566	2	probable two-compo
13	91.5	6.3	212	2	hypothetical prote
14	91.5	6.3	664	2	hypothetical prote
15	91	6.3	681	2	hypothetical prote
16	91	6.3	1217	2	2,4-dienoyl-CoA re
17	90.5	6.3	814	2	cathearin-15 precu
18	90.5	6.3	1329	2	conserved hypothet
19	90	6.2	660	2	probable inhibitor
20	89.5	6.2	886	2	hypothetical prote
21	88.5	6.1	210	2	hypothetical prote
22	88.5	6.1	278	2	fas ligand - rat
23	88.5	6.1	281	2	conserved hypothet
24	88.5	6.1	998	2	probable two-compo
25	88	6.1	531	2	conserved hypothet
26	87.5	6.0	492	2	hypothetical prote
27	87	6.0	197	1	tumor necrosis fac
28	87	6.0	204	1	lymphotoxin - bovi
29	87	6.0	445	2	hypothetical prote

30	87	6.0	2124	2	H83357	probable non-ribos
31	86.5	6.0	399	2	F83633	hypothetical prote
32	86.5	6.0	439	2	B70629	hypothetical prote
33	86.5	6.0	755	2	B75346	probable competenc
34	86.5	6.0	762	2	E98121	hypothetical prote
35	86.5	6.0	764	1	S14113	1-phosphatidylinos
36	86.5	6.0	810	2	D95256	Arp-dependent Clp
37	86.5	6.0	887	2	AC0535	ClpB-like protein
38	86	6.0	139	2	B69953	hypothetical prote
39	86	6.0	565	2	G98331	probable oligopept
40	86	6.0	568	2	E83325	probable chemotaxi
41	86	6.0	777	2	A87309	hypothetical prote
42	86	6.0	839	2	F75518	hypothetical prote
43	86	6.0	933	1	B48349	glycoprotein B pre
44	85.5	5.9	563	2	C98227	hypothetical prote
45	85.5	5.9	563	2	AB3059	hypothetical prote

ALIGNMENTS

RESULT 1
138427
4-1BB ligand - human
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C:Accession: 138427
R:Alderson, M.R.; Smith, C.A.; Tough, T.W.; Davis-Smith, T.; Armitage, R.J.; Falk, B.
Eur. J. Immunol. 24, 2219-2227, 1994
A:Title: Molecular and biological characterization of human 4-1BB and its ligand.
A:Reference number: 138426; MUID:94374434; PMID:8088337
A:Accession: 138427
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-254 <RES>
A:Cross-references: EMBL:003398; NID:9571322; PIDN:AAAS3134.1; PID:9571323

Query Match 7.4%; Score 106.5; DB 2; Length 254;
Best Local Similarity 26.5%; Pred. No. 0.32;
Matches 74; Conservative 31; Mismatches 93; Indels 81; Gaps 13;

QY	32	PPAPMAARRSGRRGRGEPGALLVPLALGGLLAVVSL-GSRASL-SAGE	89
DB	16	PPAP-----RARCVRPL-WALVAGLLLLLAACAVFLACPMAVAGARPCGSAS	67
QY	90	PAQBELVAEEDDPSELINQTEESQDPAPFLNRLVRRRSAAKGRTRARRAIAHYEVH	149
DB	68	PRLRE-----GPRLSP-----DDPAGLLDL-----RQGMFAQLVAQNVLL-	102
QY	150	PRPGDGAQAGVGVSGMEE--ARINSSPLRYNRQIGEFYTRAGLYLYCOVHFDE	206
DB	103	-----IDGPLSWISDPGLAGVSLTGLSTKEDTKELVAKAGYVYFQ-----	146
QY	207	GKAVYLKLDLVDG-----VLALRCLEEFSSATAASSLGPQLRLCOVSG-----	249
DB	147	-----LELRVVAGSGSSVSLALHQLPLRSAAGAAALATVLDLPASSEARNSAFRGQ	201
QY	250	-LLALRPSSSLRIRTLPMHLKAAPFL-----TYGELFOV 283	
DB	202	RLHLTSAGCRIGLVHLTERRAHAMQLTOGATVGLFRV 240	

RESULT 2
S27224
N-methyl-D-aspartate receptor epsilon-4 chain - mouse
C:Species: Mus musculus (house mouse)
C:Date: 25-Feb-1994 #sequence_revision 01-Sep-1995 #text_change 17-Mar-1999
C:Accession: S27224
R:Iikeda, K.; Nagasawa, M.; Mori, H.; Araki, K.; Sakimura, K.; Watanabe, M.; Inoue, Y.
FEBS Lett. 313, 34-38, 1992
A:Title: Cloning and expression of the epsilon-4 subunit of the NMDA receptor channel
A:Reference number: S27224; MUID:93050214; PMID:1385220
A:Accession: S27224

[illegible]

```

RESULT 13
A70611 hypotheical protein Rv1219c - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: A70611
R:Cole, S.T.; Brosch, R.; Partholl, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, K.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G. A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome A:Reference number: A70500; MUID:98295987; PMID:9654230
A:Accession: A70611
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-212 <COL>
A:Cross-references: GB:293777; GB:AL123456; MID:g3261726; PIDN:CAB07841.1; PID:e1299832.
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv1219c

Query March 6.3% Score 91.5; DB 2; Length 212;
Best Local Similarity 24.5%; Pred.No.4.1;
Matches 53; Conservative 26; Mismatches 78; Indels 59; Gaps 9;

OY 46 GRRCGPTALLVPLATIGL---ALACIGLLAVVSLGRASLSAOPEAEIVAE-- 99
||| ||| |::| |::| |::| |::| |::| |::| |::| |::| |::|
D1 21 GRHG-----FCGLRAIAEAGVSAAVLVIHHGSNEG---RKACDDFAEEIR 66

OY 100 -----DPPSELNPOTESQDPAPFLNRIVRRRSAPKGRKTRARRAIAHYEHR 151
||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
D6 67 SSKAAALKSNPPTWLIQMAEIESYAPLMAYLVSRSMOGSELAKKMLQKMI----- 117

OY 152 PGCGAGCAQCVCYTSGWEARINSSPLRYNRQICEFI-VTRAGILYLYXCQVHPD----- 205
||| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
D1 118 --DNAEYLD----EGVRAGTVKPSRDPRARRAFIAITGGGGFLLYLQIMHENPTDLR 168

OY 206 -----EGKAVYLLKLDLVDGVLAIRCL-EEPSATA 234
||| |::| |::| |::| |::| |::| |::| |::| |::| |::|
D1 169 AALRDYAHMDVLPISLEVTEGLADRAMYEAFLEA 204

RESULT 14
D83231 hypothetical protein PA3305 (imported) - Pseudomonas aeruginosa (strain PAO1)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: D83231

```

R:Stover, C.K.; Pham, X.Q.; Errin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.;
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
A:Reference number: A82950; MUID:2043737; PMID:10984043
A:Accession: D83231
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-664 <STO>
A:Cross-references: GB:AE004753; GB:AE004091; NID:99949433; PIDN:AAG06693.1; GSPDB:GN
A:Experimental source: strain PA01
A:Genetics:
A:Gene: PA3305

```

Query March 25.8%; Score 91.5; DB 2; Length 664;
Best Local Similarity 25.8%; Pred. No. 15;
Matches 77; Conservative 28; Mismatches 117; Indels 77; Gaps 12;

OY 28 ROAOPAPMAAR-----SOR-----RRRRGEGPTALLVPLAIGLALACGLLL 74
      ||| : ||| ||| ||| |||
Db 194 ROPOGGLLNLKRIKIVEADQORDHAMFEGEGRRRACGALLSLDLSL-----LRT 245
      ||| : ||| ||| ||| |||
OY 75 AVVSLGSRASISAOEPAGE-----LVAEEDDOPSELNPTFESQOPAPPLNLRPRR 128
      ||| : ||| ||| ||| |||
Db 246 ARGVARQPARISEEERVERKMLALSALEGTDPASQALRELAQYA-----VEPQW 299
      ||| : ||| ||| ||| |||
OY 129 SAPRKRTARRARAIAAHYEVHPRPGODGAQAGVDGTVSGMEEARINSSSPLRYNR--OIG 186
      ||| : ||| ||| ||| |||
Db 300 SNDQ-RYLLTRCISVLLKAVN-----AEKGMRAVASGEVEGRVSGAGTSLMWRDQMA 351
      ||| : ||| ||| ||| |||
OY 187 EPIYTRAL-----XYLY-----CQVHDEBKAVYIKDLVDQVLA 223
      ||| : ||| ||| ||| |||
Db 352 LFYGRSRLALLGSLVYVITAMPAPASGAMLLAAVOSLEFARNDAVAIGISFLRGIVYA 411
      ||| : ||| ||| ||| |||
OY 224 LRCEEFSATTAASSLGPOLRLCOVSG--LLALRPGSSLRITPLPMHLKAAPFLTYEGL 280
      ||| : ||| ||| ||| |||
Db 412 I-----PAAMLVSGMLLPQNNGFPLLCLLGVPLFFATLGMAYVVTAGTATSFAL 461
      ||| : ||| ||| ||| |||

RESULT 15
H83044
2,4-dienoyl-CoA reductase FadH2 PA4814 [imported] - Pseudomonas aeruginosa (strain PA
C:Species: Pseudomonas aeruginosa
C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
C:Accession: H83044
R:Stover, C.K.; Pham, X.O.; Errin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; L
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; L
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa
A:Reference number: AB2950; MUID:20437337; PMID:10984043
A:Accession: H83044
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-681 <STO>
A:Cross-references: GB:AE004894; GB:AE004091; MID:99591076; PIDN:AAG08199.1; GSPDB:GN
A:Experimental source: strain PA01
C:Genetics:
A:Gene: fadH2; PA4814
C:Superfamily: Methylotrophilus methylotrophus W3A1 trimethylamine dehydrogenase

Query Match 6.3%; Score 91; DB 2; Length 681;
Best Local Similarity 26.3%; Pred. No. 17;
Matches 79; Conservative 31; Mismatches 86; Indels 104; Gaps 20;

OY 6 FEISARRLPRLRSIGSRDGAQOAPAP-----MAARRSQRRGR----- 47
      ||| : ||| ||| ||| |||
Db 442 FRRVRLERLGVDRLRGR-----VRGELDGGQDDVYVATGIDPKRRPRIDIGGPTVLSYVD 497
      ||| : ||| ||| ||| |||
OY 48 --RGEP-GTALLVPLAIGLALACGLLLAAVSLGSRASISAOEPAGEELVAEEDDPS 104
      ||| : ||| ||| ||| |||
Db 498 VLRGAPVGAARVAIVAGAGIGEDVA--ATLVAPPSG-----OPRALGEMVLAEMGVLD 548
      ||| : ||| ||| ||| |||

```

```

Oy 105 ELNQTESODRPFILNLYRRRSRKAQKRTARNRILAAHVEHNRQSDGAQAGVDGT 164
Db 549 --NSOPGGLRPAP-----TTPAQQV-----WLLQRRPGARGAQL--CK 583
Oy 165 VSGMEARINSSPLRYN--RQIEFTYTRAGLYLYCOVHPRDGKAVUKLD-----TL 217
Db 584 TSGWYH-----RAHNRHNAVMLG-----GVETL-----KIDE--KGLLRVUGVENMLE 626
Oy 218 VDGVL-----ALRCLEFESATAASSLGPQLRL-----COVSGLL-----ALRPGSSLRIR 262
Db 627 VDNVYICAGQEPRLRQJROTSES-----LRPILIGAVAYAGELDKRAIRISGAMLAAR 660

```

Search completed: March 31, 2003, 07:59:21
Job time : 34.4558 secs

STIC-Biotech/Ch mLib

89 055

From: Schnizer, Richard
Sent: Friday, March 14, 2003 6:45 AM
To: STIC-Biotech/ChemLib
Subject: 09/245,198

RECEIVED
MAR 14 2003
STIC

Please search the commercial and interference databases for SEQ ID NOS: 1-4 of 09/245,198.

Please perform an oligo search on SEQ ID NO:4.

Please also search for nucleic acids that can encode the amino acid sequences of SEQ ID NOS:2 or 4.

Thank you-

Richard Schnizer, Ph.D.
Patent Examiner
Art Unit 1635
CM1 12E17
703-306-5441
Mail Box CM1 11E12

Point of Contact:
Thomas G. Larson, Ph.D.
703-308-7309
CM1, Rm. 6 B 01

Searcher: _____
Phone: _____
Location: _____
Date Picked Up: 3/24
Date Completed: 4/1
Searcher Prep/Review: 15
Clerical: _____
Online time: 20

TYPE OF SEARCH: 4
NA Sequences: _____
AA Sequences: 3
Structures: _____
Bibliographic: _____
Litigation: _____
Full text: _____
Patent Family: _____
Other: _____

VENDOR/COST (where applic.)
STN: _____
DIALOG: _____
Questel/Orbit: _____
DRLink: _____
Lexis/Nexis: _____
Sequence Sys.: ABSSOT
WWW/Internet: _____
Other (specify): _____

THIS PAGE BLANK (USPTO)